Transcript of Cyberseminar

 Session Date: 3/11/2015

Series: HERC

Session: Econometrics Introduction and Identification

*This is an unedited transcript of this session. As such, it may contain omissions or errors due to sound quality or misinterpretation. For clarification or verification of any points in the transcript, please refer to the audio version posted at www.hsrd.research.va.gov/cyberseminars/catalog-archive.cfm or contact: susan.zickmund@va.gov*

Heidi: Today's presenter is Dr. Susan Zickmund. She is the Director at the Qualitative, Evaluation And Stakeholder Engagement Core, and Core Director at the CHERP Methodology Core at the VA Pittsburgh Healthcare System; and Associate Professor of Medicine and Clinical & Translational Science Director at the University of Pittsburgh. Let me make you the presenter here, Susan. We can get things started.

Susan Zickmund: Okay, thank you very much. Can you see my slides?

Heidi: We can see your screen, yes.

Susan Zickmund: Okay, excellent, great. Alright well, welcome everyone again. As Heidi had shared, I am Susan Zickmund here at the VA Pittsburgh Healthcare System and in the Center for Health Equity Research and Promotion, and CHERP. My goal is to talk about traditional qualitative research as well as large scale qualitative research.

We will go ahead and begin. I have really two main goals for this methodology seminar. The first is I would really like to talk about what is traditional qualitative research. My goal in this is to some extent to give a standalone and fairly brief half an hour long methodology lecture talking about the basics or the ABCs of qualitative data collection and analysis in the health sciences; so, very much specific to HSR&D investigators.

Then in the second half, I would like to talk about really sort of plowing new ground and discuss the approach to large scale qualitative methods that I have been using within a VA study known as the Disparities and Satisfaction with Care or DISC, an HSR&D study. Let me start off by sharing sort of the organization of the seminar. First of all, I would like to start off that principle investigators – that is defined as anyone on the call – can use to help conceptualize qualitative and mixed methods research? Then I will transition into talking about what is unique both in the real opportunities and the challenges of using large scale qualitative data.

In terms of part one of these questions, my goal is to give the most simple of heuristics that will help you; or, if you are talking to anybody who is thinking about adding qualitative research. Just thinking through the steps that are needed to make certain that the data collection and the analytic part, it has been thoroughly thought through. If you want to envision this as if these slides did not exist anymore. We certainly hope that is not going to be the case. You could just go through these very simple steps and think through what did I need under why? What did I need under when? As well as just to give you some information that might even function as a checklist. Say yes, that is right. I thought about that. I thought about that. That is what I can pretty much do in terms of traditional qualitative in 30 minutes time.

Let us go ahead and begin. First of all I would like to start with a bit of a word of definitions. Because I know those can get to be confusing. I purposely use the term qualitative in the title because I am focusing on different approaches to textual methods, a.k.a., qualitative research. But I will sometimes use the language of mixed methods. I just wanted to explain what the difference between the two happens to be. All mixed methods research involves qualitative data by definition. Mixed methods involves utilizing multiple methods, so quantitative and qualitative. It can be in the form of intervention trials and in depth interviews. It could be a chart review and in depth focus groups for example. But the goal of mixed methods is to somehow intentionally integrate qualitative and quantitative data together.

I will sometimes use the term mixed methods because it makes sense given the examples I am using. I would draw your attention to the web link, which is the best practices for mixed methods research in the health sciences, which is coming out of the NIH. It is an excellent introductory guide and sort of the gold standard for qualitative rather for mixed methods. Okay, so in terms of yourself. I would like you to think about if you are envisioning yourself using qualitative to answer the question why. Why does qualitative make sense for the project that you are interested in?

Having a clear rational, I think is important. Now, I am a methodologist. I work as, also an investigator. But I wear a methodological hat often. I work with individuals who come to me and are interested in qualitative research. I will sometimes ask why. I usually will ask why. I will receive an answer sometimes. Well, the reviewers told me to do it. Or, they will say it is, for example, PCORI. The Patients Centered Outcome Research Institute. I work on a lot of those particular proposals.

Qualitative is very typical in those. People will say it is PCORI. I have to. Now that is a good reason to start the conversation. But I think it is always important to make certain that you think about what it is about your question that can be best aided through the use of qualitative data. Also, I think it is important to make sure that whether it is a grant or an article, that is very clearly articulated.

Then I will talk a little bit more about different layers of data. But make sure that every time you choose these qualitative data, you also justify that. But the things to keep in mind was that very first question. Moving to the next, when; so, the principle investigator again. Anybody who is thinking about putting a proposal together. I do encourage people to work with qualitative experts just as I would encourage folks to work with the statistical expert. I always, I am very grateful for the insights that I gain from my statistician when I work with him on projects – to think about the entire project. Determine when qualitative data collection, should an analysis, should it occur? It is thinking about it, a classic mixed method design. Here inherently, there are qualitative and quantitative data together, which makes it the mixed method.

There are two main ones, and again drawing upon the NIH guidelines. There is convergence – convergent design; which is the defined as to merge concurrent qualitative and quantitative data together to address study aims. At the end of the talk in the second half, I will give you an example of a convergent design. Then there is a sequential design to have one data set built upon the results than the other. When you think about when, I will share with you that the vast majority of mixed method studies, of where there are many; and increasingly many in health services, HSR&D. They are typically sequential. There are different types of data collected within a mixed method study.

They start off qualitative and then move to surveys. You may have a chart review. You may end with focus groups. A PI needs to decide when during a particular year and during a phase data collection can occur. Things just to keep in mind when you are envisioning that. Let us imagine you are going to be putting an intervention study together. It may be that it would be helpful to have a process of tailoring or even creating that intervention with very specific qualitative data. Maybe you want it at the end of the intervention to really sort of drill down on the experiences of important stakeholders and patients. Maybe the clinical individuals who participated as well or \_\_\_\_\_ [00:07:34] to deliver a \_\_\_\_\_ [00:07:35] intervention. Maybe you might be interested in prior to the dissemination, just thinking through when it might make the most sense for your project. All projects are different. There is never necessarily rights and wrongs.

Okay, moving to the third question, what? This involves a type of data collection. There is a lot of different what's out there. There is ethnographic work where you observe a certain clinical situation that you want to learn more about. You could do a content analysis to try to understand the persuasiveness of a media campaign. In health services, oftentimes the data types of collection that we use come down to interviews and focus groups. Now there are whole Cyberseminars on both of these and look at entire classes. What I would like to do is just share a few rules of thumb that you as the principle investigator can think about in terms of coming up with the best type of data, the type of data collection that you want to use. Interviews typically produce the best data. I think it is my view. But the literature, I think will back that up as well. It would end up individual uniquely tailored to capture the specific information that you need for your project.

Focus groups can be perfect at a certain situation. Also recognize one of the redeeming qualities of a focus group is that typically they require less time, less effort, and less resources. If there are fewer transcripts and overall less analytic work. If you have five focus groups with let us say ten people each. You have 50 versus 50 interviews, the amount of data collection time, the amount of transcription and analytic work, it is going to far less from focus groups. But those are some aspects to think about.

Let me share some others as well. Shifting back to interviews, they are typically considered the gold standard in qualitative research and health services research. You can have these amazing interactions. They are one-on-one. You can really drill down to the very specific questions you are looking to get answers for, and know that person is likely to answer, unlike focus groups. Now one of the things is that if distance is an issues; so, for example, if you are doing multi-site studies. Your participants are dispersed. Then there is certainly also telephone based approaches to data collection.

I could share that and this is sort of a transformation over the years of my views. I now believe that telephone interviews are every bit as good in terms of the data that you collect. I think in terms of participant burden, and also being able to make sure that you have the widest and most even selection of participants in your study so that you do not bias it toward the people who are near you. I think it is the best way to approach interview data. There is also less administrative structure that is needed, though I have mentioned certainly, their analytic costs are higher. You compare that to focus groups, which have still remained face to face until our technology becomes better. They need to be face to face sessions.

One of the really amazing things about focus groups is they are dynamic. They capture the interaction between the participants. They allow everyone including the principle investigator to walk away with more information than any one participant has in the beginning. That is something that is really special. That is why focus groups can be perfect when the rationale is strong. They do require just to be – as you are thinking through what types of data collection you want to have. They do require support at the site. You need to schedule a room. You need to deliver food. For multi-site studies where there is no RA support, they can be a challenge.

I have worked recently on some PCORI applications that are related to large pragmatic trials with sometimes up to 20 or 30 sites. Many of them more rural sites where there is no research support; and trying to have focus groups there would be difficult. Therefore, there might be an example where you might say well, if we'd do better to have targeted interviews using telephone all across the various sites with some kind of sampling matrix. Also, there are times when you have individuals who are so busy or because participant burden would be so difficult for them to get to one place. Sometimes then a focus group would not be the best choice. Providers could be hard for example; unless, of course, you find that the providers are actually and have like a lunch time or a certain period of time when they are all together.

Now, I did mention layers of data collection. Increasingly, I think inspired by the stakeholder works that has emerged from PCORI. I am increasingly seeing within QUERI graphs and within HSR&D merit reviews. There is more focus on layers of collection. Different types of data collection could occur at different times within a single study. For example, it may be that an early phase stakeholder engagement and where your goal may be to tailor or to design an intervention.

It really may be best served by having focus groups. Because it may be that here is what our intervention looks like. How can we prove it? What do you all think? Have you had these experiences? Having that group dynamic can be the perfect way to start off an individual project depending on what that project is. Then you may choose to move to later phases. They want you to experience the intervention. I really wanted to ask you about very specific variances and facilitators so that we can understand how to revise it and make it an even strong intervention.

Think about it as various different layers of data collection. Now something I always want to talk about because it is important to investigators to think about. It can be a little bit mysterious. That is the idea of thematic saturation; which is in essence a sample sized calculation for qualitative research. Now, qualitative research always had come to medicine from the social sciences. Let us use an example of an anthropologist who might get a three-year doctoral grant or a fellowship to go someplace and do data collection.

That person, their salary is covered. They are very concentrated on the data. They do not have to necessarily worry about how many interviews do I have to participate in, in advance. They would do interviews until they are done. They could do field notes and think about an interview being more. Then when they start to hear the same sorts of things again and again, they can say well, this part is done. I can now move on. It is more complicated for those of us in health sciences where we more often rely on grants. Because we really need to have an idea of how to project when thematic saturation might occur.

One of the things you need to really think about is I use the language of cell, which is you need to think about the amount of interviews that you need for the scope of your work. It is in part\_\_\_\_\_ [00:14:49] by how homogenous your population may be. If you have a single cell, so you say I want to know what medicine fellows think about their training? Let us say if you have a one site. These are all of the folks. They have been trained the same way. We really want to know what their thoughts are. It is likely given that is a pretty homogenous population.

You may be able to achieve that saturation with that single cell, that single group between 20 and 40. That would be my guess. I would say 25 is my guess of what I think probably it would take based on that description. Now, let us imagine you are interested in gender differences. Or, you want to look at differences in experiences for people who are local coming from that university or coming from outside. Then if you had 25, then you might be looking at 50, somewhere between 40 and 50. Because you want to look at a very specific difference.

Our goal is to look at gender differences. Every time that you multiply the cell; and every time the population becomes more diverse, you need to think about your sample size. In the third bullet point, I talk about the fact that here is an example of something where we start to have a very diffused population of racial differences in urban versus rural diabetes patients. There could be a lot of different factors. You need to think about what is a reasonable cell based on the amount of groups that you have. If you have two versus you have 40, your sample size per group is going to get, if you can – allow it to become smaller for the 40s. Because you are still picking up the saturation across all of the 40 groups versus two. It is still a relatively small. You may double it for two. But when you get to 40, you might bring it down to…. The smallest group that I have ever seen in the literature.

That is, my experience as a qualitative researcher for a couple of decades now is that I do not think you can really go below ten and nine. I think if you really, even if you have a gigantic study you need about the saturation about at that level. Hopefully this demystifies the saturation and sample size with the qualitative event. We can also talk about that in the question and answers, if it would be helpful. Now, let us move to the question of how? There are specific approaches to qualitative analyses. Again, there are whole courses. I know Cyberseminars have been done on them as well. Let me just share in this notion of checklists. Let and make sure you think about qualitative methodology, and codebook development, coding approach inter-coder reliability, and thematic analysis.

Moving to that, qualitative methodology may again be the most confusing aspect in part because you need to rely on somebody who has spent probably years studying some of these methodologies even if it is a mentee, mentor relationship or a member of your team. I thought it was interesting a few years ago when we had HSR&D, the national convention. The interest group of mixed methods research came together to try to give suggestions on what to – guidelines to give people who were viewing qualitative abstract. We wanted to come up with a series. The one that everybody – there was like universal agreement on was that there need to be some type of qualitative methodology. Because it is the underlying philosophy and the sort of theoretical approach that you would take to the entire project. It is also important to make sure that you are guided by someone who understands qualitative work much like you would expect that with a biostatistician.

Biostatisticians have often as far as I understand, reasonably similar training, particularly if they are\_\_\_\_\_ [00:18:32] they are coming out of the health sciences. Qualitative methodologists have different approaches all of which can be absolutely legitimate. But it is important that if somebody who has spent some time thinking about this to help guide a project. Again, you do not necessarily have to be the qualitative methodologists; but know that you have one who could give you an idea on your methodology. Having an accepted approach is important. Now, also you want to think about the specific one.

In health services, the most typical one that we have seen is grounded theory, often described as a modified grounded theory approach emphasizing grounding the information in the data. In addition, ethnography is the way that anthropologists observe and take field notes. Personally, and I know I have said this before on various workshops. I think Crabtree and Miller, particularly their book Doing Qualitative Research, they are coming out of the area of family medicine. They have various ways that the methodologies that they articulate in the intro chapter to doing qualitative research. I typically focus on their 1992 edition. But it is, all of their approaches I think are really excellent. Because they come out of – that they, out of the clinical reality that we in medicine and indeed health services experience.

I always recommend people to taking a look at Crabtree and Miller. One of their approaches is called Editing. It is a grounded theory\_\_\_\_\_ [00:20:04] approach, really focusing on the iterative, open ended way of making sure that you allow categories to emerge out of the data. Another one is known as Template, which is where that process is occurring. Then you may take a codebook and apply it to a new set of texts typically with a refinement period. Then also something that is really a fascinating approach called a quasi-statistical approach that really helps us try to understand the power that qualitative can bring not only in terms of, to invoke the situation, but also that there are frequency issues that can come in contact with some of the other more quantitative data.

That is what all I have highlighted. Because I will focus more on that for today's talk and particularly in the second half. Now another part that I find is that it is very difficult. I think the hardest thing in qualitative data is to do a codebook. I also find that the degree of it feeling mystical as well. I have a few slides just to emphasize that with a codebook as PI, and thinking about the process. You want to make sure that you when your grant, you write about. In your process, you focus on the codebook development. I have divided this into three stages.

The first, open coding, this is where you have transcripts. You are reading, and reading, and reading everything that comes to you and sort of noting things that you think are interesting and surprising. Things that are specific to your research questions that you have. Then let us imagine you have an interview study that has 100 interviews. Twenty interviews or 20 percent of the cases that you have, let me emphasize here of your study. You would probably go through about 20 of those interview. Then start listening and noting things that you want to capture. How you want to capture. Just writing down things, codes as we call them; just specific words that are helpful for you as a tag that would help you be able to capture that information.

I work with a lot of fellows, and postdocs, and medical students. Oftentimes, they have very short, very small studies. They might be interviewing ten people because that is all they can do. Then sometimes you will need your entire study, if it is reasonably small. It is sort of a sliding scale enough to be able to really understand the topics that topics are going to come up frequently. That are important and that are surprising; never underestimate the importance of surprise. Then the second process is an ongoing one. It is an ongoing loop for\_\_\_\_\_ [00:22:32] the next cases that you have available. Or go back and read the ones that you have; and just continuing what we call memoing or writing notes to yourself, to your team members, anyone you are working with of what you think is it important.

Refining, typically I encourage of keeping all codes unless you are going to be splitting them because a code is too big. Or you are going to be bringing codes together. I have happy. I have gleeful. I looked at the back later. I thought those are the same thing. Then I am going to be bringing them together. Note which ones remain constant across the various cases. Does this codebook seem like it works for the next one, and the next one? What in grounded theory is called the constant comparison approach, which is endemic or is shared by all qualitative approaches? What themes do we hear that are coming up that are new? What is important?

You keep refining this, refining, and going through this stuff until you find yourself with a fairly stable codebook. Then continue if you can. Then before I get to the third step, when you finally feel that you are ready. You have worked with your team. Other people have looked at the other transcripts. They agree with you. You have done what you need. You have gone through – let us imagine that you have gone through 50 of your 100 cases.

What you want to do is when you are ready to start, you… I am going to encourage you to have a co-coder. You get together with the other person you are coding. You say now we are official. What I think is important if you are – you finish with the 50th, you start fresh with 51. You go on up to 100. Then you go back to one. The goal is not to start your coding by going backwards to 49, and 48 where you will have already applied that largely that codebook already and discussed it. Then you also want to make sure you have a very key and precise definition of your codes, the dates they were created. What is known as the audit trail. Provide inclusion and exclusion criteria for codes; give a clear and borderline cases. As much detail on your codebook as possible; and recognizing that you will also need to have decision rules about when to use a particular code. When another one might be appropriate. Or, when in fact you feel that both codes are actually needed.

A lot of this is\_\_\_\_\_ [00:24:53] in this process. When you have gone through that, you will end up having a codebook. As a principle investigator, you also need to think about when to select – who to select as your coder. You need to decide what is your total number of coders? I joke to my students like and to pick any number as long as that number is two. If you want your life to continue to be as happy and as joyful as it was as you started this project. I have been on a lot of coding teams in my life.

There is a different philosophy which is the next bullet point. I will mention that in a second. But I have been on teams of two. I have been on teams of three. Well, we are all trying to do adjudication, even kappa statistics. But it is really difficult to have three. I think – and certainly one, you do not have the insight of a second set of eyes; which is really very helpful. There is the notion of traditional coders. I have been on teams. It is a very traditional approach. Outside of the health sciences but also especially in the health sciences to have your team be members of the\_\_\_\_\_ [00:25:56] team. Then you need to understand who is going to look at this document? Who is the co-coder? Who does it shift to and develop the \_\_\_\_\_ [00:26:04] rules? Do you feel that you as the PI should be a coder? Can you be objective? Or, is it an intervention.

You want to say I really want everyone to like it? Then you might want to – that is a simple way. Let someone just to help address some of the bias issues that sometimes can come with qualitative coding. You also need, and you think you about your proposal. In reality, thinking about your training process, if you have outside coders or people who are not already bringing those skills to the table. Or, they do not have the skills for your project. How are you going to talk about your project in relationship to their skills? Then think about the determining inter-coder reliability. Well, you could have a single coder.

Anthropologists traditionally do not code with anyone. It is typically expected in the health sciences within the HSR&D's community. Are you going to have two within adjudication process? Would you like to have two coders? I always recommend it in an adjudication process, going through each code even if you agree on it to make sure you know what you are agreeing on. Talking about what you disagree; then you can also add a kappa statistic on, particularly if it is a big enough project. If you have less than and this is just a number that I think seems to make sense, but less than 20 cases. Or, if you have like three focus groups, it may be difficult, if you do a kappa statistic.

But some type of adjudication process. Kappa are very sensitive to the size of the study. I always try to include this line. Because Landis & Kock, or as I have found are the only statistically oriented investigators who have ever really explained how to interpret what a kappa means. Then in the teams that I am on, we usually try to shoot for certainly about 0.70, which obviously would fall into that substantial category. Then it is always helpful to have a computer data management system. It is up to you to decide which one. I am going to explain a new system that we built for ourselves and for our DISC study.

Traditionally there are two. But there are other ones as well. Atlast.ti and Nudist are the most popular ones. We store your interviews, focus groups, transcripts, the codebooks you have. Certainly it is going to help capture your codes; and also in the back side of your codes are going to be your quotations. Software enabled a level of management complexity and textual complexity. It is not possible with notes at all alone. Because there is a lot of paper and a lot of keeping track of things in qualitative research. It is giving you a principle screenshot of what Atlas looks like.

We are not going to go into Atlas at this current moment. But just to share with you what it looks like. I wanted to emphasize that any computer program you use is for management. It does not actually do the analysis. It is always the excellence of the team and the coders that do the analysis. However, it is very helpful to be able to retrieve that information.

Then the last slide on this first half of the talk. Thematic analysis is the most typical way to code to capture the major themes related to the research question. This involves reading verbatim quotations categorized under a specific code. That is, of course, part of the coding process. You may find that you are going to be doing a thematic analysis that will involve a single code. Let us say transportation barriers are so rich, you are going to write an entire article of that. Or maybe a combination of codes and that is why sometimes there is a distinction between a code and something larger where you have stepped back. You call it a theme. Typically in qualitative research, you will have an article that will focus on three to five major themes. There can be a lot of different derivations. But I would say that is sort of a classic qualitative thematic analysis.

That is, ABCs of qualitative in half an hour. I would like to share with you just some of the development of an approach to a large scale qualitative study based on my DISC study. Now why do this? Something has really\_\_\_\_\_ [00:30:18], it is fascinating to me. It is this idea of – and I work with a lot of individuals who do secondary data analyses. Well, when you have big data, your ability to make generalizable claims, and to see patterns, which is I think is really important in the healthcare system is – it is amazing. I find it very attractive.

However, as a qualitative researcher, I really want to maintain the depth of the qualitative insights. I have always been interested. Is there a way to harness the strength of both of them? This is my humblest account at answering that question positively. I can share with you a little bit about that experience. Our DISC study focuses on disparities in satisfaction with VA care where we found based on the patient satisfaction data that in the past that it has come from the VA. There have been discrepancies indicating that African-Americans, Hispanics, as well as women may not be as satisfied on all of the difference healthcare domains. The more dominate group; which is white males in the VA. We were really interested in trying to understand why.

In our large mixed methods study, we decided to target 103 – 1,350 from 25 VA Medical Centers using the telephone, of course. All data came from the University of Pittsburgh as you know, the VA Pittsburgh. We had a one hour interview that included the surveys. I share on this slide. Our focus was on males and females. Half of that sample are males and half are females. We have this one-third African-American, one-third white, and one-third Hispanic. Then again, we collected from 25 VA Medical Centers. I will tell you a little bit those centers in a moment. Then, we had six different spells per VA Medical Center, women and men. Then the three race ethnic groups together; which meant that we had nine participants. That was my choice. The lowest number that I have found; and I feel that is correct. Of course, thematic saturation which gave us our total of 1,350.

Let me tell you a little bit about our sites and our interview script domain. We have 25 DISC sites. Most of them were chosen because of high African-American, or Hispanic, or moderately high numbers of both. Then the HSR&D reviewers encouraged us to include less predominately, or I should say predominately white serving Veterans. I am sharing with you to say yes, a visual representation of, we basically ended up for the most part having major sites given that we were interested in diverse and minority serving\_\_\_\_\_ [00:33:07] and that were at the geographically, at the border of the country with the four non-white minorities I have not placed on that, but shared among that five.

Then in addition, we had a – we drew our conceptual model for the DISC based on a meta-analysis of the patient satisfaction literature, and patient satisfaction, and patient experience. Some of these domains focused on the entire cohorts such as the overall care. The most important is domain. Because we asked people about their experiences with the VA care. Most of what they have to say comes out in an unprompted fashion; which helps us to know what they really thought before we start to drill down. Then often we will hear those same themes come again. Then there is some themes that – or rather domains that would only effect the subgroup; the specialist in mental health, and let us say women's health.

Those are the sort of the main domains in the DISC script. I will share a little bit about our sample. When the DISC interview study, we mailed out over 8,000 invitations. I will bring it down to the individuals who are eligible to complete the interview. That was close to 2,000. Now there has always been; and though there are some movements affront on this. They needed each site for a form – a written form that needed to be sent and signed, and received based on IRB requirements to audio recorded voice. After that consent document, you moved on to 1,386 with the ability to complete 1,207 to date for the DISC script.

It is a study. Our responding characteristics, it is stratified samples. Even distributions of females and males, white, African-Americans, and Hispanics. We are interested in doing a small thematic qualitative study on the other category. The age distribution, fairly reflective of the VA in total. Then we also found that VA only users, it was excluding the dual users. It was certainly more than half the population at 62.4 percent. Again, I do not want to get into the, any of the actual data of this yet. I just want to share with you our approach and now talking about our inter-coder reliability and our codebook. This was the largest team I think I have worked with.

Certainly the teams\_\_\_\_\_ [00:35:39] complexity of codebook. We had needed seven coders for this process. Developing a codebook required input from all coders. I began as I traditionally have. I had like two master coders with the most experience listen over a good period of time to 200 interviews; memoing and developing themes. We sat down. We talked about it. We recognized things were slightly different as it is the domain of say axis versus respect, and humaneness. Then we brought the larger team together who had been doing a lot of transcriptions. They knew a good deal. They knew themes that were coming up. They had not developed the codebook.

Then there was an interesting process of discovering that it took a lot. Basically I learned that you need all of your coders to be at the table in building the codebook. That was new to me. That was one of those, the surprising finding. It took us a very long time from August to November with daily meetings. I think I counted that I had 700 hours with them. I know they had much more in trying to build this, the codebook. Also, not only because of the complexity of the size of particularly given the fact that the first question is well tell us about your experiences with your healthcare.

The breadth of that first question was more than anything that I had experienced before. We needed to be able to have a very broad series of categories; but then can allow each one of them to very in depth to be able to effectively capture the kind of information. Of course, we are asking people about their experiences; so their satisfaction, and dissatisfaction. I will tell you a little bit about the codebook. But let me mention the reliability process who had seven coders operating in pairs; and doing inter-coder reliability across basically 3.5 teams.

The inter-coder reliability started with us doing two individuals, 20 interviews. Then they would shift to another team. We discovered that the more that you could bring that down to five interviews as we shift more frequently in our different rounds, the better. Our goal is to have a single mind. I use maybe a timely given the nature of what we have heard about Leonard Nemoy, and this sort of Vulcan mind meld where everybody is on the same page on the codebook. Just shifting them with each other and doing an adjudication process as well as continuing throughout the course of the entire project even once the inter-coder reliability is done; having large team meetings to make sure that we avoid the coding drift. To maintain consistency and discuss new themes that might start to emerge.

That has been our process. I will tell you it is just a little bit more about how we capture our story. In the DISC interview script across each one of our domains, we have this…. There are three things that are identical. There may be additional questions. But three things that each one – that frames each one of the domains. We start with a five point Likert scale ranging from very satisfied to very dissatisfied. Can you share with us your thoughts about your overall satisfaction with your VA care from very – this is very satisfied all of the way down?

Then the subsequent open-ended question. What contributed to your rating of your overall satisfaction once – your satisfaction with your overall VA care? Then we might probe other questions or not depending on how rich that particular domain needed to be. Then we would end each one with a question. What could the VA do to improve your satisfaction with that particular domain? With the goal that as the Veterans have expressed their views that we want to understand what they think might be a good idea in terms of solving it. That is how we framed each and every one of the domains.

Looking across and listening to the stories that Veterans shared, we came up with these main codes. There is satisfaction, dissatisfaction, improvement. There is structured and then hierarchical; and as much as possible, parallel structure. The characteristics, I really like my doctor. The clerk was kind of rude.\_\_\_\_\_ [00:39:57], the way that humans talk about other humans. Communication, interpersonal communication, and their views on their quality of care. A lot of questions, and a lot – the biggest domain that we have is on access, continuity, and coordination of care.

There are thoughts on the brick and mortar facility as well as their local CBOC. Their attitudes on cost, non-medical services parking, for example. Yes, we have a code called Starbucks and pharmacy. Those are our major ones. At least they do drill down from there. Between satisfaction, and dissatisfaction, and improvement; again, that's a parallel structure. There are 500 codes roughly that we have for each. Then we use that for each one of the domain names.

One of the things that is complicated and difficult for qualitative researchers is that particularly relying on a software like Atlas.ti, you have codes. Atlas has its limits. Because you have to be able to see your codebook. There is a little window for it. If somebody talked about it. I loved my care. But one time I went to the ER, and I mean, nobody knew what they were doing. I had a nightmarish experience. Then I had this one doctor, I feel it was neglect. Somebody is angry. They are talking about it.

Classically, you would have maybe awful experience, nightmarish experience, whatever code would capture that kind of sentiment. You might have an ER code. You might have having let us say an incompetent, an incompetence code. You then might have a provider code. There is no way to really bring them all together.

We worked with our data manager, David Obronsky here in CHERP. He developed a computer system for us that enables us to have a drop down menu. If we have a code let us say, a bad experience in the ER. I am making that one up. He can associate a code that lets us determine who did it involve? More generically, so I am not going to talk about doctor in the ER, but doctor. There is also just moving to some of our other categories, and I do not want to overwhelm. Just putting what is on a, nicely put on a screen; but categories to capture service. That where each time there is a default of unspecified.

That is probably what often occurs. But we could then go to the ER, the Emergency Room, and click that. Even because we are interested in our 25 DISC sites that we are capturing organizational information and characterizing. We are also interested in trying to understand are they talking about that site? Are they talking about their CBOC? Or, could it be this is an entirely different VA? We can also, all together be able to know that story is – those series of codes are linked to the various ways that I described. Then that, for me personally, the first time I had an opportunity to be able to experience this. For example, a quote, one of our quotes from a woman saying the whole aspect of being a woman Veteran.

For example A, and this is the CBOC. It is kind of funny. But they have a little room that says "Women Only, Veteran's Healthcare". The whole thing about that came because of MST, Military Sexual Trauma. The staff goes by. They are left, and right, whatever. I have even left suggestions in their box. You go by there. I would not sit in there. There are men in there. It is kind of pointless. They are not aware. We have a code under our facility area called women do not have own space. This example would have that code.

We also have a global code. Every time someone talks about MST in any context with other people and their own experiences, they will receive that code. They will have in terms of the employee, in this case it is unspecified. We know they are employees. But we do not specifically know what they are doing. Services I guess, unspecified; but we do know the location is CBOC. The computer program will allow us to capture all of these details related to this specific quotation. There is a lot here on this computer.

I just want to walk us through a few things. If a coder sits down, and they would type in the idea of the participants. They would type in their ID. They would also look at the various domains. If you were able to see in the middle that the number two overall VA care. That is that main VA care code, it has been – like what domain has been clicked. It automatically brings up whatever questions or specific interview questions underneath that domain. Right now, somebody is coding what contributed to your rating of overall VA healthcare.

The whole codebook is available. You could time stamp it. You could put the quotation in. You could put notes about this is a really good quote because it says this and this. Also, with auto-fill, it allows the coder to be able to know what the cod is. Then be able to mention who it involves. Where it occurred and then also something about the location. This kind of computer program is where we found ourselves going to help us capture not only the size of our data but also the breadth of our codebook.

Then just a little bit more information. It is hierarchical coding structure. The codes are stored in a number of variables; which is helpful. Because we decided on a structure that was the same for each domain. The same name might be there, but a different…. The coders know because they know what screen they are in. But we can analytically know that is overall versus access. There is also upward and downward collapsing so that on the example on the screen here, we might just be interested in appointment scheduling and timing. General is that specific code. That is what it always means; appointment scheduling timing in the data set. But maybe eventually we want to know something about time to get to an appointment. That allows us whether it is something we are going to report or being able to find specific quotations. It gives us a lot of dexterity of upward and downward mobility in the coding structure.

There is a lot to say about this. It is a very exciting study. I would like to share just a few things before I transition to the end. What can we learn from large scale and qualitative? I think there is a lot more that we can. I will come back when we have our complex models that we will bring in our survey data, our health utilization data, and the demographic information as well as all of these codes. But other things that we can certainly understand that we can gain from them. For example, a sample size that is this size allows us to look at – I will give you an example. The ratio satisfaction and dissatisfaction themes per Likert scale.

We have ultimately, we really do not know what it means to be somewhat satisfied versus being somewhat dissatisfied. I can guess based on having heard hundreds of these examples. It may be that someone is somewhat dissatisfied when the only issue is access. But if they have problems with their provider in terms of interpersonal or confidence, I think it is much more likely to move it. But we can know as opposed to anecdotally. Or, for example, with a computer system like this, we can actually drill down to understand, if we have a code.

I am just giving an example, nightmarish. My ability to see what is employee type or a service, it might be associated with that, or even superlative, the outstanding codes. From a healthcare systems perspective, there is a lot of information that we can gain from this. That is what I will share for now. I want to do – I want to share a sense of appreciation for my co-investigators. I am just mentioning the ones here at Pittsburgh and also the research assistants, and the photos that I have here. Because it truly does take a village. Everyone has just been so excellent on this project. Alright, so I am going to open it up for questions. I look forward to answering them.

Heidi: We have received a few questions in for the audience. If you do have a question, please send those in using the questions pane on the GoToWebinar dashboard on the right-hand side of your screen. I will just start from the top here. We will work through with it. The first question does it have to be a person on the phone? Can interactive voice response systems yield useful telephone surveys?

Susan Zickmund: Yeah. It is interesting. There is a company that is an offshoot of Carnegie Mellon University who has that. We did talk to them more about recruitment and seeing if they could get, and use these systems for clinical trials. My understanding and my sense is that this is a pretty advanced company in that software. I mean, it is pretty good on the robotics. It is that it is not there to the point where it might be able to technically to provide it.

Also, I would say that I spend so much time training the interviewers to have an empathetic and humane reaction. If they do not, if they are icy, you do not get very good data. I think the idea that people are willing to talk to a computer at this point, that I think is the issue that is going to make me say no. I think that if we still…. I do not want to talk to a computer. I do not like it when they end up on my answering machine. I would say no. I do not think we are.

Heidi: Okay, great, thank you. The next question, does the dependent variable determine the number of cells?

Susan Zickmund: That is not a qualitative person asking that question. That is fascinating. The answer may be yes. But I want to – again, you are talking to another qualitative person here. You want to understand your focus. I am going to make a distinction between the focus and a purposeful sampling. Because it is something I needed to skip over. There was a degree of remorse because there is those degree of complexity. If your focus is gender in whatever you have collected. You are interested in \_\_\_\_\_ [00:50:01] training. We will go back to the medical training and whether or not there is differences. How they perceive their training base with gender. You want to understand that you have enough folks to be able to comment on gender. I think that probably is your dependent variable.

I will let the audience mentally say yes or not to the answer of that. Purposeful sampling, which is the\_\_\_\_\_ [00:50:24] sampling structure where you try to get the…. This is various purposeful sampling. But let me say you went back to the variation. You wanted to get people who are different. You want people who are happy. They are not so happy as well as people in the middle. Then you will get a lot of variation there as well. But that is not what I am talking about in terms of this idea of doubling or making sure at the cell level.

There you have a broad perspective. Then what I think you need to think about is that you have a less homogenous population. Your overall "cell," it needs to be bigger. I want to make sure that I do not just have medical students from let us say the University of Pittsburgh where I am at, or fellows. I want to include people who are coming across the country. Maybe they are coming from areas that are outside of the U.S. They are all medical. They are all fellows.

There are some overlaps in their trainings. But it is a broad one. I really want to make sure that I have all of those groups having an opportunity to share their insights. Then I probably would not have 25 or 30; I know it is not that homogenous. I am likely to have maybe 50. It depends on how broad and how much they share things. That is a long answer I just want to feel that purposeful sampling versus that I was talking about for thematic saturation, it needs to be – it needs to be a little bit clearer than I\_\_\_\_\_ [00:51:44].

Heidi: Great, thank you. The next not a question, but kind of a question. I tried finding the Crabtree and Miller book resource online unsuccessfully.

Susan Zickmund: I too have – every once in a while, I will lend mine out. I rue the day when I do that. Because it is also pretty expensive. It is not there. It is a wonderful item to purchase and to not lend out. Yes, you are right.

Heidi: Thank you. The next question here. Do you have any experience or thoughts on online focus groups:

Susan Zickmund: Unfortunately, I have experience with telephone focus groups. These have not been my ideas. They sometimes, particularly when you work with trainees. You need to make things work. I know that those do not work well. Because you cannot see people to be able to determine. In a Cyberseminar I once gave on focus groups, I believe, and I researched it at that time for the Cyberseminar. I think that is going to be much more effective in the future.

The technology needs to be able to allow us to see each person, and to be able to see them, and hear them. We need more specification than what I have experienced using the system that I have had so far. I do not think you would want to see people in the background and not know who\_\_\_\_\_ [00:53:14] talking about. I do think we are going to move in that direction. As a qualitative researcher, my feelings on participant burden have become – they have really transformed over the last ten to 15 years. I think it is important to use distance as much as possible. We do not have individuals spending the day on the bus. They can come in to do a face to face interview or to do a focus group. I think ethically they are important. I do think technologically we will get there. I have not seen it as of yet.

Heidi: Great, thank you. The next question, how can one join your team?

Susan Zickmund: I do not know. I do not quite fully know what it means to say join my team. You can e-mail me. I am the only Zickmund in Outlook. That is a claim I feel like I can safely make.

Heidi: Great, thank you. We received two different questions in about different software or analysis programs. I am going to kind of combine these two. If you are familiar with Dedoose or Nvivo, I think is hopefully how they are pronounced?

Susan Zickmund: Dedoose, no, that may be a newer one. Nvivo is very popular. It is used by a lot of individuals. Again I run a methodology research course where a lot of people come with data half done. They come to our Core. We do use Atlas.ti. I mean, one of the reasons is that I – in the '90s, I had a kind of 16-week course in Atlas 4.2. I have come to know it well.

It was fairly obscure at that time. I think the fact that it is so popular is an indication of it is a very user friendly. Its presuppositions are, or do not dictate your own analytic approach. I have really appreciated that. I am sure that Nvivo is also fine. It is just something that I have known, one. I have trained everyone to use that particular program. We are increasingly trying to see how we can work with investigators who are using Nvivo so we can bring…. It just a burden on the investigator if they have to shift things over to Atlas because we have done it. I think dexterity in the computer programs are important. Obviously, this new program for this study is something that we have developed.

Heidi: Great, thank you. We just got a comment in here that VA does not support Dedoose because it is web based data storage. But it is great for mixed method research.

Susan Zickmund: I thought that. Well, it is great to know. I had that for\_\_\_\_\_ [00:55:56].

Heidi: The next question we have here. Do you usually set limits on the number of codes that you have to make the coding process more manageable?

Susan Zickmund: When in the past – I mean, I had a study at the University of Iowa where I had 2,500 interviews. It was a giant study. My codebook was four pages. It was fairly easy to bring a lot of new coders in because we could train them. I had, in the past, when I used Atlas, I wanted to be aware of the fact that I had Atlas. You have to look at your codebook in a window. Now I do a lot of scrolling. I personally, and I have talked about this in other Cyberseminars, do questions specific. Or in this case domain specific; and it is both a question and specific coding. Then we have sort of long codebook.

I do not set limits. Particularly now, I am not setting limits. I think you need to capture the data in the level of details that the study dictates. Our DISC study, because it is so large, and because there is a large diversity; we feel this was the codebook that we needed to have the tools to effectively answer in. I think that is more important. If I am working with somebody who is new. They are going to take an analytics part. I really encourage them to have precise but succinct codebook. Because it is just much more easy for them to use. It is not a full answer. I think you need to come up with a codebook that is most appropriate for the project you have, the resources, the knowledge that you have the focus\_\_\_\_\_ [00:57:31].

Heidi: Great, thank you. The next question here following the question about sampling. With the focus of the research on gender and race, do you care about heterogeneity and the type of medical condition and type of services they have received?

Susan Zickmund: Well, I mean, we care. But it is not an inclusion, exclusion criteria. What we ask in terms of our inclusion criteria is we ask that they have had an outpatient visit in the last 12 months so that they have some experience at one of those 25 sites. Then we also have a – when we get the demographics section, part of that demographic. We have the self-reported Charlson Comorbidity Index. Then we also collect health utilization. We have collected ICD-9.

We will take the information that folks have shared. We will be able to characterize it. But we did not want. Our goal is to understand in general. The population in the VA particularly with the folks that have raised ethnicity and gender. What are their attitudes? We did not want to make it more restricted. Because then the generalizability, if we dare use that language\_\_\_\_\_ [00:58:43] the qualitative. It starts to be diminished.

Heidi: Great, thank you. We have three questions left here. Let us, if we can try to get through those quickly and wrap things up. The first question here how do you deal with recall bias with focus groups? What is too long? I assume it also depends on what the topic is. I am thinking specifically about ICU patients and families.

Susan Zickmund: Recall bias, and whether it is focus groups or interviews is an important issue. Sometimes not even recall bias, but the ability to recall. If you need people to mention very specific technical information, then it would be – it would behoove one to have – to contact them more quickly. I think what is important is to try to make sure you are more reasonably even across whoever you are talking to.

If you ask people about their experiences in the ICU, then you make sure that if you cannot bring people to the table for a focus group until three months out – everybody who has been three months out. Then you can see what that experience or roughly. It is what it is. But I think, the more precise in our design, the more that we can control it that way.

Heidi: Great, thank you. Next we have can you please comment on qualitative analysis of written responses to open-ended questions? For example, embedded in a survey rather than interactive interviews or focus groups.

Susan Zickmund: Sure. I have done this\_\_\_\_\_ [01:00:15] an analysis with Brian\_\_\_\_\_ [01:00:17] on health literacy amongst high school students. In your methods, you simply describe what it is that you have done. I think one needs to understand that the level of complexity of different kinds of data is going to be more or less complex. But the qualitative approach in terms of codebook instruction, precision of analysis. All of that will be the same, it is just probably going to be sparse. It is certainly more than a written, an open communication would be, verbally.

Heidi: The last question we have here. Do we have natural language processing of these focus groups or survey transcripts?

Susan Zickmund: I will presume that the question is do we have the capacity to use natural language processing? I mean, it is something that I am really interested in. I have been part of what is known as the CHIR, the Consortium for Healthcare Informatics Research. It is a bioinformatic people in the VA. I think there is a lot of overlap, not only analytically in terms of what bioinformatic people do and qualitative folks. But I think that these are tools that can go together in a way. It is an area that I am fascinating; and I would like to move in that direction.

When I look because I wanted to do a piece that was specifically on the use of bioinformatics and qualitative, I did not find work currently moving in that direction. I am working with Scott Duvall, and\_\_\_\_\_ [01:01:52], and his team on trying to see if we can use bioinformatic tools for qualitative research. I think it is something that we will get and move in that direction. But I know very little about it that is not published to date.

Heidi: Great, thank you. That wraps up all of the questions we have. I know we are just past the top of the hour here. I wanted to thank the audience for hanging with us, if you guys could hang on for just another minute or two. When I close the meeting out, you are going to be prompted with the feedback form. If you could take just a few moments to fill that out. We really do read through all of your feedback. It allows us to make some changes to the programs dependent on what you all are looking for. Dr. Zickmund, I really wanted to thank you again for taking the time to prepare and present for today's session. We very much appreciate all of the work you have put into it.

Susan Zickmund: \_\_\_\_\_ [01:02:41]. Thanks everyone.

Heidi: Thank you, everyone for joining us today. We hope to see you at a future HSR&D Cyberseminar. Thank you.

[END OF TAPE]