Session date: 9/12/2016

Series: VDM

Session title: Practical Guide to the CDW Mental Health Domain

Presenter: Eric Hawkins

 *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.*

Moderator: So joining us today we have Dr. Eric Hawkins; he’s an investigator at the HSR&D Seattle Center of Innovation for veteran centered and value driven care. He’s also the associate director of the Seattle Center of Excellence Substance Abuse Treatment and Education and associate professor in Department of Psychiatry and Behavioral Sciences at the University of Washington. Joining him today is Carol Malta; she is a research health scientist specialist also at SCESATE at the VA Puget Sound Healthcare System. So at this time I would like to turn it over to you Dr. Hawkins.

Dr. Hawkins: Thank you Molly thanks for that really nice introduction. And I also just wanted to remind everyone that there will be a few points throughout this presentation where we’ll pause and ask for questions. But I also wanted to invite those of you who may have questions after this presentation to please feel free to give us an email. I think our contact information will be at the end of this slide deck. Let’s see, it’s not advancing now.

[Informal Background]

Dr. Hawkins: So there are several objectives to this presentation. I’m going to start off by describing what the mental health domain is and where the data come from and then Carol will describe the types of data that are available in the mental health domain. It will include some of the measures, the available data elements. And then an emphasis on some of the data extraction issues. Carol will also describe how the data can be used with particular focus on some of the limitations and challenges that are associated with using data in the mental health domain. And then I’ll end our presentation with an example of one of our recent research projects that use the audit C which is from the mental health domain, the corporate data warehouse.

Moderator: So for our audience members today we would like to understand what best describes your role as a user of the CDW Mental Health Domain; so go ahead and click the white circle next to your response. Those answer options are researcher, clinician, administrator or policymaker, quality improvement or other and that would include non-BA. And if you are selecting other, please note that at the end of the presentation I will put up a feedback survey with more extensive list of job titles; so you might find your exact title there too marked down.

It looks like we’ve got a very responsive audience, already 75% response rate; so we appreciate that. I’m going to go ahead and close out the poll and share those results. So the overwhelming majority 59% of our audience are researchers, 3% clinicians, 4% administrators or policymakers, 13% quality improvement and 20% selected other. Thank you for that and we’re back on your slides.

Dr. Hawkins: Thank you. And so the information that we’ll provide in this presentation is really relevant to three groups of users; so research, operational and clinician users. I think it’s important to sort of briefly not that access to these data largely depends on how it will be used. So for example, research users as most of you are probably well aware you’re given permission to use these data on a project by project basis whereas in kind of operational users typically have full access to a mental health domain. And then finally clinicians are usually given some sort of limited access to the data. Often times it will be limited to their \_\_\_\_\_ [00:04:07]. Another poll question.

Moderator: Thank you. So for the next poll question, have you ever used the CDW mental health domain? Yes or no? And it looks like we’ve got people responding, about two-thirds of our audience so far. We’re right around 80% response rate so I’m going to go ahead and close the poll and share those results. We have 29% of our respondents saying yes they have and 71% reporting no; so thank you for those responses.

Dr. Hawkins: So great. The majority are researchers but it sounds like the majority have not used the mental health domain; so hopefully there will be some valuable insight in here that we can share. So the mental health domain contains information that’s related to mental health assessments and individual assessment instruments and these are instruments that are typically completed as part of the standard clinical care. They’re collected via the mental health assistant which will describe a little bit – in a little bit more detail in a subsequent slide. And there’s some timing issues with these data that are probably important to keep in mind. Vinci documentation indicates that information didn’t really start to get into the mental health domain until about October 2007 and that was only by some of the four regions. It really wasn’t until March of 2011 that all four regions contributed data to the mental health domain and we’ll mention this last point on several occasions throughout this presentation.

But a lot of you are probably familiar with the fact that many of the measures that we’ll talk about or that are available in the mental health domain were being used prior to the mental health assistance and domains origin. And in that case we’re not going to talk much about how to do this but you can – these data are available, it can be extracted from the corporate data warehouse, not from the mental health domain and typically you’ll need to use health factors, which are these little data tags that you help you identify where these data might be.

So on this screen shot and this gives you an example if you were to go into the mental health assisted long data – this would be the first screen that you would see. The one sort of unique thing about mental health assistant software is that it’s actually embedded in CPRS. And it takes two log ins to get into it. You actually have to log in to CPRS and then you’re going to have to log in to the mental assistant software.

So as I mentioned it’s in CPRS, it’s released in 2002 and it’s essentially a graphic user interface for the mental health testing package. It can provide secure computer environment for patient administered testing and then the other thing to note is information is also collected via clinical reminders and clinical reminder dialogue templates. And we’ll talk a little bit more about how that is done. Essentially those are instruments for those of you who are familiar with some of the annual screens that occur in the VA, alcohol misuse, depression and PTSD. That information is typically completed and enters the mental health domain through clinical reminders and clinical reminder dialogue templates.

So on this screen shot it gives you really the second screen that you would see in the mental health assistance administrators what it’s referred to and this is really where a clinician will go in and sort of select or identify the measures that they’d like to use. They’re essentially put into a queue, you can’t see from this screenshot but on the bottom right hand corner there’s a button that you can click that essentially says “begin to administer the measures”. They’ll go through each measure one by one and each question one by one. And then this is just a screen shot to show you what a clinical reminder actually looks like. This is the national clinical reminder for annual screening of alcohol misuse. Just something to note there’s essentially two windows on this screen, one that’s sort of imposed on the other. The screen with the first two items of the audit C. If you were to go into the mental office assistant and try to administer the audit C after you started that process this screen right here would like identical to what is seen in this clinical reminder. And that’s because this clinical reminder actually links to the mental health assistant.

I think before I turn it over to Carol to begin describing the types of data that are in the mental health domain this is sort of a place to pause and see if there are questions?

Moderator: There are no questions at the present.

Dr. Hawkins: All right, so Carol you let me know as I advance these slides if you need me to advance them.

Carol Malta: Okay thanks Eric. Now we’re going to shift from talking about how the data are collected to looking at how they’re stored in the CDW. So I’m going to walk you through the layout of the mental health domain and some of the mental health data elements and then take you throw some of the lessons we learned in working with these data.

So on this slide here what you see is a simplified version of the relationship diagram that you might be familiar with from the CDW share point. And what I want to draw your attention to is for this domain there are four main tables and those are on here in blue. And then there are eight domain tables and those are on here in purple. The dash lines let you know which tables are directly joined to one another. And there’s a brief summary of the types of data you’re going to find in each table on the slide here.

I’m going to walk through a number of these tables in detail and go through some examples on how to join the data, but I really encourage you to use the link here that’s in the upper left hand corner to go to that relationship diagram on the CDW just to get familiar with what all is there and the overall layout.

I’m going to jump in by looking at the survey table, which is over on your left hand side of the screen. So if you go to the next slide what you see the output if you run a very simple query just on that survey table it gives you a list of the available measures in the mental health domain. \_\_\_\_\_ [00:11:08] draw your attention to when you run this simple survey you get over 200 unique measures, but note that on some of these there are duplicates – not really duplicates but multiple versions of the same measure. So for example if you were really interested in the back depression inventory, so you go down on the output here to row 28 and you see okay there’s an entry for the back depression inventory. But then you look one row below and you see oh, well there’s also the back depression inventory second edition and then below the back depression scale. So before you really get going with anything you’re going to want to sort out which of these am I interested in or is it some combination.

The other thing to note is that when I ran a quick query on this about a month ago I came up with 200 measures but only 97 of them had any kind of administration data. And then if we take that a step further going to the next slide what you can see is that even those that are in use may not be used very frequently or by all the sites. So just to kind of orient you to this slide, what you see on the output here the first two columns are going to identify the particular measure for you. The third column let’s you know the total number of sites that we’re using that measure in the past year. And then the last column is going to give you the total administrations in the past year. So if you or someone, say a researcher coming to this and you really wanted nice national data on any of those measures in the first half of this table you’re going to be out of luck because as you can see they’ve just not been administered very frequently.

In contrast if you look down to last three entries here, these are the measures that Eric mentioned primary care PTSD screen, the patient health questionnaire to and then the alcohol use disorders identification test \_\_\_\_\_ [00:13:08] items and these are \_\_\_\_\_ [00:13:13] measures and you can see in the past year well over a million administrations, all the sites are using them.

I just want to contrast that briefly to the brief addiction launcher which is highlighted in blue in the center of your screen. So what you see here is actually four different versions in use of this particular measure and not all the sites are using any given one. But it may be that if I can combine those in some way I’ll get good national coverage; those are just things to consider as you’re working through the data.

Now if you move to the next slide what I’m highlighting here is that you’ll also want to think about timing; so your availability of particular measures is going to depend on what that measure is; so those things like those three national screen \_\_\_\_\_ [00:14:07] the patient health questionnaire in that primary care PTSD screen. You can see we have really good data going all the way back to 2008, so over a million administrations of each in any given year; so very good coverage. And then I wanted to contrast that to a lesser used measures, so the PTSD checklist. There’s a civilian version which is the C that you see in line four and then the military version which is the next line down. So you can see that’s a much smaller cohort of patients who are being administered those measures, but again if you go back all the way to 2008 we have consistent administration; so good data availability on those.

Looking at the three measures below that, so lines six through eight on your screen there you can see that those measures didn’t really come onto the mental health assistant until 2012; so you can’t get data on those until a later date. And then I wanted to also bring in the brief addiction monitor revised just to let you know that if we look at line seven and eight here we see the brief addiction monitor is still in use and used very frequently but the brief addiction monitor revised came on board in 2014 and in 2015 roughly 10% of your administration were using that revised version. So if I am going into this data and I don’t include those I’m really missing out on a big chunk of people so I need to keep those things in mind when I go to pull the data.

So the next slide is just sort of summing up and talking about these are some things that you want to think about when you’re first approaching the data. You want to think about why were the data being collected, were they associated with performance measures or metrics. If they were you’re going to tend to get national data. How were they collected, was it systematic, was it a part of the national clinical reminder such as those that Eric showed you, or was it more of a local initiative. And then you want to think about the timing of the data is important. So did VA quality initiatives influence the data collection, did any changes occur to the data collection over time and that could be a change in the patient cohort who was being administered the measure or perhaps a change to the measure itself.

So from here I kind of want to dive in more deeply to some of the main tables in the mental health domain. So I’m going to start with a survey administration table and this table is generally critical for most queries because it will let you join up with other tables in the mental health domain but also lets you join to other domains in the CDW. So for example if you want to know something about where the survey was administered you can use your location SID to link up to the location table. And if you wanted to know more about who was administering these measures you can use that order by SID and the administered by SID to link up to your staff tables; so you can get a lot more information. And then of course your patient SID will let you link up to your patient demographics and characteristics.

So expanding outwards if we go to the next slide you can pull in the survey results table as well and what you’ll see here is that this will provide you with a score for your measure as well, which typically the field that you’re going to typically be interested in is the raw score; for most of the measures that I pull the transform scores are blank. It may just be the nature of the things I’m looking at there may be other measures where those are actually filled in but I typically am just seeing the raw score.

So you may say to yourself “well this looks good, I can use the survey administration data to get my patient identifier” and when the survey was administered and then the survey result will give me my score and that’s all I need, I’m done. But I would really recommend taking it a step further and pulling in your item level responses. And to do that you’re going to be relying on the remaining two tables the survey answer table and then the survey answer text table. And we’re going to go through some examples to show you why you might want to do that; so moving to the next slide.

To go through this example I’m going to be using data from the audit C and so I just want to take a few minutes to introduce this measure for those of you who aren’t familiar with it. So the audit C is made up of the three consumption items from a larger 10 item alcohol assessment that was developed by the World Health Organization. So the full audit also includes items on drinking behaviors and alcohol related problems. Just note that the audit C and the audit are stored in the mental health domain; so if you’re really just interested in those audit C results you’ll want to make sure that you’re screening out any of those full audit results.

So what prior research has found is that those three consumption items that make up the audit C are a valid screen for alcohol misuse. And the next slide will just let you know what those items are. So the first question is asking about frequency; so how often did you have a drink containing alcohol in the past year? So this is just getting at any alcohol use. The next asks about quantity; so how many drinks did you have on a typical day when you were drinking in the past year and then the last question asks about heavy drinking episodes, so how often did you have six or more drinks on one occasion in the past year. All of these are scored from 0-4 so we expect a total score on the audit C between 0 and 12.

Within the VA the audit C has been used as a national screener for alcohol misuse since 2004. However between 2004 and 2008 the audit C often wasn’t completed on non-drinkers and as Eric mentioned this was during the time when these data were typically collected via health factors. And so sometimes you will find a health factor indicated that someone is a non-drinker but it will be sort of a separate health factor. It won’t have the audit C tag with it so you’ll have to kind of pull multiple things together to get a full picture of what was going on during that period of time. However if you move out to 2008 forward all patients were asked question one on the audit C which is frequency of any use and then the questions two and three get auto-filled for those folks who are not drinkers. And as Eric mentioned all of this is collected via clinical reminder and next slide just gives you a snapshot of that that Eric already showed you. But that means that show data collection for this particular measure is much more systematic and clean than some other measures might be.

If we go to the next slide what you’ll see here is if you just run a very simple query and wanting to know the raw scores on the audit C and you pulled this from the survey result table you’ll really hope that you’re going to get results that range from 0-12 because that’s what the range on the audit C is. But what you actually come up with is a range of raw scores that go from 0-57. And so really this is just to let you know that you can get some noise when you rely on that raw score. This appears to be more of an issue with some of the older measures that have been in the mental health assistant from the very beginning. And that my guess its caused by some incorrect questions or response options getting tagged, particular measures at certain sites, but just know again, this is just something to be aware of that you’re going to want to look closely here and make sure you’re not getting these out of range responses.

The other issue that can come up with scoring though is you can have a case where the raw score looks just fine; it’s in that range of 0-12. But if you went and looked at the item responses you would see oh well you know this really shouldn’t have been scored at all due to some response error; so I want you run you through some examples of that. So with the audit C the first row on this table what you’ll see is somebody who had a raw score of zero and this is pulled again from that survey result table. But then if you go in there and pull in their item responses to the questions on the audit C and again just to orient you really quickly the question one is frequency of any use, question two is quantity and question three is frequency of six or more drinks.

What you’ll see in that first row is what we hope to see if someone says “I never had anything to drink in the last year” then those next two items aren’t asked and that is perfectly okay because we don’t expect them to answer about quantity if they didn’t drink anything in the last year. However if you go down to the next entry you see a raw score of four and this is someone who said “I drink four or more times a week” but then you also have these questions two and questions three being skipped. And in that case I can’t really rely on the raw score because of someone who really only answered one item out of three item questionnaire and really based on their response to question one should have answered those other two items. The next two rows just give other variations on that and rows three and four you see in each of those one items was skipped. So again, for me I’m not going to use those because if it’s only a three item questionnaire I really want to see everything filled out for me to trust that raw score. The last two rows on here are a little trickier because on those you have all the items filled in but the responses don’t necessarily follow from one another So for example you have somebody who said, “Oh I drink for or more times a week” and “I drank seven to nine drinks typically on any occasion, but I never drank – I never binge drank; I never drank more than six or more drinks on one occasion”. So that kind of question okay can that make sense? And the same with that final row somebody who said, “Oh I didn’t drink anything over the last year” but then went ahead and answered the next two questions on quantity and whether they binge drank.

So you’d really need to consider well for my purposes do I want to use these data or not? And my whole point with this is really to just say that that raw score doesn’t give you a full picture necessarily of what’s going on. So this slide gives my overall recommendations, so given that you know the reliability of those raw scores is going to vary by your measure, your date of administration and your site. You’re really going to want to check them out thoroughly first before you rely on them and just to remember that those older measures will tend to have a little more noise.

So what I always do myself is I pull in the response to individual items on the measure and then I calculate the total score on my own. So to do that I rely on the survey answer table and I link it up with my question text and my multiple choice answer text and then my free answer text and we’re going to run through an example of how to do that.

To go through that example I’m going to be using data from the brief addiction monitor or BAM and it just is a bit of background. The BAM was initially implemented in VA in 2012 and it’s designed to be used in addiction specialty care programs who are encouraged but not required to administer at baseline when someone first comes into the clinic and then again at follow-up. So what we’d expect to see with our BAM data is that there’s going to be a lot fewer administrations than the audit C just given that it’s a smaller patient cohort, but we’d also expect to see multiple administrations per patient over a short period of time. Also just to note that the BAM was really designed to facilitate measurement based care and given initiatives in VA right now to increase measurement based care when you may see administrations of this measure increasing in future years.

Just a little bit more about the measure. It’s 17 item questionnaire that was developed by VA investigators in Philadelphia. It assesses areas relative to substance use and also to recovery. Often it’s the responses to individual items on the BAM rather than a total score that are of interest. And for this measure I’m typically producing it for administrative reasons or for clinics to monitor themselves and see how they’re doing; so they’re typically again more interested in those individual items. And also I’ve already noted that there’s four different versions of the BAM. The original BAM which is still very much in use and then the revised version of the BAM is becoming more common. And then the last two versions are used pretty infrequently so I’m not really going to go into them here.

But typically what I’ll do is I’ll emerge my BAM data with my revised BAM data and the fact that one of them is – the original BAM is all categorical and then the revised BAM has categorical and continuous responses makes this a little bit more complicated; so I’m going to go through how to do that.

The next slide just gives you a snapshot of what the revised BAM looks like in the mental health assistant; so you can see here on item one there’s those categorical response options for the clinician to select and then the remaining items on there are all continuous responses that people fill in.

Within the CDW, if you go to the next slide you can find information on the response type for individual question from that survey question table that was one of our domain tables and you’ll be looking in the field survey response type, which is over here in the third column. And so what we can see here is on the BAM we have three different types of response type, again this is that revised BAM. You’ll have a date field and then you’ll have multiple choice fields and then also into \_\_\_\_\_ [00:29:21] fields. So again, important to know that because of where those data are going to be stored.

So if we go to the next slide. Here is just a summary of where you’re going to want to pull your data from; so to get your responses to multiple choice questions you’re going to want to join your survey answer table up to your survey choice table. And then to get all other response types whether it be data integer, free text you’re going to want to join your survey answer table up to your survey answer text table; so it can get kind of confusing. I provide a \_\_\_\_\_ [00:30:02] just for your reference but what I really want to do is walk you through the output of what you’d get if you ran a query like this. So if we go to the next slide what we’ll see here is in that first column it’s giving you the question text for every question on the BAM. The second column gives you the response type for your reference. Normally I wouldn’t pull this in; it’s more just to show you what’s going on. The next column, the third column gives you the survey choice text and this is all pulled from the survey choice table. And then the final column gives you the survey answer text that’s pulled in from the survey answer text table.

So what you’ll want to notice here is that for your multiple choice questions you’ll see that the survey answer text is filled in with a response but your survey answer text is null. And so you’ll see that all the way through here. So for example in the second row the question is in the past 30 days how would you say your physical health has been, and this is a multiple choice question. And the person has answered excellent. And you can retrieve that from the survey choice table and in the survey answer text table you’d find that that response was null. And then for all of your integer or any other actually non-multiple choice response you’d find that your survey choice text is going to be listed as missing and your survey answer text will be filled in with a response. And it should be the right time for \_\_\_\_\_ [00:31:49] response so for example if it’s a date, it should be in a date format. And if it’s an integer it should be in an integer format. You can get occasionally some weirdness in there but that’s normally what you’ll see.

So for example on question – on the third line here in the past 30 days how many nights did you have trouble falling asleep? We get missing data in the survey choice text but we get a zero filled in in our survey answer text; so that’s where we’re retrieving that particular response from. So this just gives you a picture of where these different data are stored within the mental health domain.

I know I’ve covered a lot of information so I just want to take a second to pause here and see if there are any questions?

Moderator: We do have quite a few questions. Starting with where does PHQ2 live? Does it have date and time stamps?

Carol Malta: Basically if you look back at – I don’t know if you want to go back to the slide that gives the administration over time. Yeah so what you see here is the patient health questionnaire, two is in the second row of this table; so you can see yes there’s consistent data on that measure going back to 2008; so quite a few administrations and all of those administrations will have the data that we’ve talked about. They will have the – everything that’s in that survey administration table; so you’ll know when it was administered, you can see who administered it, the location where it was administered. You can link up to the raw score on it. For that particular measure it’s going to be far more relevant for you to know the actual responses to the two items. So you can also go in and pull in those item responses in the way that we talked about. So that one you actually only need to pull from the choice text table because it only has categorical responses; there’s no other data type on it.

Moderator: Okay another question, we are building a perioperative risk tool and want to pull in an audit CPHQ2, get seven PTSD to identify risk patients. What advice would you give us in terms of which score to use? The most recent average over the past few years or some other?

Carol Malta: That’s a very good question and Eric is going to talk a little bit about that in terms of research context but it will really depend on what your goal is. I mean for me given the want to know what’s going on with this person, I’m assuming when they come in the door I would probably pull in the most recent, that would make sense to me. But it may be that you have something else that you’re thinking about in terms of getting a picture of what’s going on with people. Maybe you want to know something about how they’ve changed over some period of time. So that will really be driven by what you’re intent of using those measures is. Eric I don’t know if you had anything to add on that?

Dr. Hawkins: No that makes sense to me.

Moderator: Okay here’s a related question. We’d like to identify patients with latent undiagnosed mental health disorders at the time of major surgery. We’d like to detect this by examining PHQ2 \_\_\_\_\_ [00:35:35] in the year after surgery among those with normal scores in the year prior to surgery; does this approach seem reasonable?

Carol Malta: Well it can be a challenge I would say because – well so the audit C is probably a little cleaner because you have something that again you’re just \_\_\_\_\_ [00:35:58] alcohol because you have something that has been shown in the research to actually pick up people who are you know depending on the score qualifying as having alcohol misuse or severe alcohol misuse. You have a nice cut point and you have a measure that is very frequently used in VA because every patient is supposed to be administered this measure every year.

And PHQ2 you will certainly – you only have two items so it’s not going to be as nice as having a PH29 on everyone but at least it’s something, again that’s administered on a lot of people. But keep in mind if you’re trying to use things like the BAM it’s really administered on a very select group of folks and so you’re only going to be picking up the people most likely who have started substance use treatment episode so they’ll be limitations there.

Moderator: Okay we do have quite a few more questions. I’d like to ask the presenters if they want to continue giving the content or take the questions now.

Carol Malta: I think it makes sense to move ahead and then we can answer more at the end – make sure we cover everything.

Moderator: Yes I agree.

Dr. Hawkins: All right. So for the remainder of this presentation I’m going to talk a little bit about how we use the audit C to answer a research question in one of our projects. There are actually several questions that we were trying to answer but for the purposes of this presentation we’ll just focus on this one which is what is the prevalence of alcohol misuse among OEF, OAF veterans with and without mild traumatic brain injury.

So just to give you a little bit of background rates of alcohol misuse among OEF, OAF veterans are as high as 40% with rates higher among men than women and among younger than older veterans. 10-15% of OEF, OAF veterans have sustained at least one TBI during their deployment. And these conditions are problematic because of the negative outcomes that are associated with alcohol use among those with TBI. So for example alcohol use may impair recovery from TBI, it may increase the risk of recurrent TBI as well as the development of alcohol use disorders. So for these reasons early detection intervention is important to prevent these negative outcomes. So what do we mean by alcohol misuse? So it’s the \_\_\_\_\_ [00:38:44] ranging from drinking above recommended drinking limits to meeting diagnostic criteria for alcohol use disorders.

And so when we do a research with the audit C we typically categorize audit C scores into risk groups and Carol already sort of alluded to some of these risk groups. So I’m going to just start off on the third row here with data. And you’ll see audit C scores of three to four for women and four for men are at a risk of alcohol misuse and a brief intervention at least in theory would be helpful for this group but the thing I wanted to highlight is in the VA there’s actually no action required. So it really isn’t until you get a score of a five or higher where some type of intervention is required. So for patients who have an audit C score of five to seven we typically think of them as being at risk for moderate alcohol misuse and in this case if a person was to screen with an audit C score between this range, the VA requires that a brief intervention and/or a referral to specialty treatment occur.

And then for the last row or group here is those with audit C scores of eight and higher. We typically classify this group as being a risk of severe alcohol misuse and that’s because there’s an associated, there’s an increased risk of alcohol use disorders among this group. And then typically the required action really should be either an assessment and/or referral to specialty treatment.

So this slide just gives you sort of an outline of several steps that we perform when extracting data from the \_\_\_\_\_ [00:40:34] as we – I mentioned earlier we’re going to go through the audit C for this; so we’ll talk about each one of these steps in a little bit more detail in subsequent slides. So the first step is really just to identify our study population and this case it was all OEF, OAF veterans that were enrolled in the VA with an audit C screen documented in the medical record in the year 2012. So after we had that sort of population we start to look at the data a little bit more closely and in this case we’re just looking for some outliers. It’s a bit difficult you know without seeing trends year over year. In this case you can at least – we had 12 months, so we could look at months but essentially you’re just looking for anomalies in the data and one that can occur is like a below average number of patients with audit C at a particular site for example. So if were at our site via \_\_\_\_\_ [00:41:35] and we know that there are only a handful of audit C’s for a particular month that would be sort of – it would raise a red flag to us. There is something that probably didn’t go right in the extraction process as we have to go back and look at that in a bit more detail.

The other sort of issue to consider at this point in time, and Carol mentioned this earlier is there are health factors likely to be out there and this is sort of a legacy effect in theory that really shouldn’t be anymore. But I – and I can’t – I don’t know exactly why they do exist. If it’s there are sites that are still using the old clinical reminder that existed you know prior to 2008; or if they somehow can – each site can – it’s not ideal but they can actually sort of customize the clinical reminder site. And so regardless of the reasons one question to raise is do we want to use any of the patients who are identified as non-drinkers, and in this case we chose to.

There are roughly I think 1,000 to 1,200 but we did include them. And this is going to be typically relevant just for the measures that existed prior to the mental health assistant; so as Carol did a really nice job showing that’s primarily a PH22, the primary care PTSD and the audit C. So on step three is where we really want to start to examine individual items a little bit more closely and decide how to address several issues.

So one of the issues that’s missing data and as carol mentioned if a patient – sorry if an audit C screen – if the first question is scored with a zero, even if the other two questions are missing we can score those questions two and three with a zero. Question one indicates that a patient is a non-drinker so we’ll by default assume that they weren’t – their responses to two and three would be the same.

The other issue that can occur however is if there are just response patterns, missing patterns that reflect something other than items two and three being missing and item one being never. And as Carol mentioned that’s a little bit more difficult. For our purposes we’ve always just sort of assumed that these are probably not valid administration of the audit C and so we exclude them from our analysis.

There’s another issue, once again Carol sort of touched on responses to items that are inconsistent or don’t make sense and we’ve highlighted those bottom two rows. These are a little bit more of a gray area; you’ll have to make a determination about whether or not you want to include those. For our purposes we’ve, at least to date we’ve always included them and you know obviously one of the things you’d probably want to look at is just to see to what extent this is occurring.

So moving on, still on step three. There’s a couple other issues that will occur and I don’t think we’ve talked about these yet. One of the ways that we used in determining what we think is the most appropriate response or decision is to determine or sort of try to identify how it might impact our research question. So one of the issues is there will be duplicate audit C’s and these come in two versions. One version is there might be multiple audit C’s that have the exact same administration date and all the response items are exactly the same as well. And this is the easier of the two versions to sort of deal with, essentially you’ll want to pick one of those three or multiple. Again, I don’t know exactly why this occurs, there’s lots of different scenarios that are possible, but in our case we’ll – we take one of them and then we exclude the others.

There’s also another version of this where for a single patient there are multiple \_\_\_\_\_ [00:45:37] with the same administration date but the response items and therefore the total score are actually different. And in this case this is a bit more tricky, our rationale to date has been to take the higher of the multiple – the higher audit C score of the multiple administrations. And our rationale is we have pretty good data from other research, that patients tend to under report the extent of their drinking. So if we’ve got two administrations one with a higher audit C score and another we’re assuming that the higher one is probably a more accurate reflection of their drinking patterns. And then one other issue that can occur and will is if you’re looking at a timeframe, so in this case we were looking at 2012. Remember that screening is supposed to occur annually but it doesn’t have to happen every 12 months. It can happen – the interval can be longer than 12 months but it also can be shorter.

So for example we may have had patients who actually were administered the audit C say in February of 2012 and then again in December of 2012. They’re likely both valid administrations but you’re probably going to want to pick one or the other. And again I think this depends on what your research question is and how choosing one or the other choices is best. In our case we, at least for this project I believe we actually picked the latest. And often times that will have to do with – it may have something to do with your access to data. If you just got your 2012 and you’re interested in an outcome that might occur after that audit C screen – if you are limited to just 2012 data that outcome might not actually occur until sometime in 2013. But if you use the administration that occurred in February you can – you may have sufficient time to look for that outcome. So again, your research questions will guide your decisions there.

And then step four is when we start to link the audit C data to other VA demographic and clinical measures that are located in other domains NCEW. So in this case we were interested in TBI diagnoses. So we extracted those, gender and age because of their association with increased risk of alcohol misuse. We wanted to present our data using those covariant and then we also defer the characterized sample you’re interested in other measures like marital status and race and service connection disability as well as ethnicity. And we also linked our data – audit C data to substance use and mental health diagnoses. We could have just as easily selected a medical diagnosis if we wanted to. The point of that slide is you can link data from the mental health domain to various other domains in the CEW.

And then lastly there’s really sort of two points that I want to make on this slide. First I’m going to orient you to it a little bit because it’s busy and I apologize. You’ll see the second and third column are essentially where women are represented on this table and you’ll see GDI and no TBI and the final two columns of the table are men and again TBI, no TBI. The other thing to sort of keep in mind is you’ll note that we started – this is where age sort – we were able to sort of reflect our results by age. And if you move over – I won’t belabor this point but you can see that the rates of alcohol misuse, again using our definition of an audit C score five or higher or essentially 25.4 among those with TBI and – men with TBI in 22.6 with men without TBI. And then women just as we would expect the rates are lower 8.3% for women with TBI and 7.2 for men with TBI.

The other thing just to note is you’ll see again sort of all the patterns that we expected the rates of alcohol misuse increased as veterans with younger. The other point that I wanted to sort of highlight on this slide is that you may also be interested in some of the individual items. They may have some information that’s useful. So in our case if you recall item three of the audit 3, which essentially askes how often you drink six or more drinks on one occasion in the past year, that’s sort of – we define that as a heavy episodic drinking episode and we were sort of interested in knowing what the prevalence of episodic drinking was among the sample. I’m not going to go through those results but just you can see that essentially all of the relationships and patterns that we saw in the audit C, greater than five results were seeing the same thing in heavy episodic drinking.

So a couple resources that you might find helpful is really this nice \_\_\_\_\_ [00:50:58] that will give you some really sort of useful information \_\_\_\_\_ [00:51:05] just for a couple years. We actually didn’t know it was even in existence until being asked to present the data. And then the other thing that might be helpful is there – it’s a bit data to this point but Drs. Kay Lysell and Rick Owens presented on the mental health assistance in informatics cyber seminar back in 2007. And that might actually be some additional data, at least how things occur for the structure and details associated with the mental health assistant.

And here’s our contact information, and again if you do have questions please feel free to send us your emails, we’ll try to get to them as quickly as we can. And then lastly any questions?

Moderator: Okay Molly am I – so we have quite a few questions is there a way to access a list of the measurement scales that are in NCDW?

Carol Malta: Again on – towards the beginning of my portion of the talk it gives you an example of how to do that. So they’re all contained in that domain table, \_\_\_\_\_ [00:52:30] but again to note that because the measure is there doesn’t mean there’s data on it so you really do have to do some more exploring to see whether you’re getting good coverage by it. And then I believe, but Eric you maybe can speak more to this but if you were a site and you were wanting to monitor something just because the measure shows up on this list in the CDW doesn’t necessarily mean it’s going to show up in the mental health assistant. Because I noted when you had your screen shot, for example the \_\_\_\_\_ [00:53:08] inventory was a BDI2 that was showing up. So that’s what’s going to be available to you. So just because it’s listed here it may have been kind of the – they may have said “Okay we’re not going to use this anymore, we’re using this” and so that’s what’s going to show up for your use in the mental health assistant.

Dr. Hawkins: Yeah and that’s correct. The list is sort of fluid. I actually went in and counted in the mental health assistant and just sort of counted the number of measures that were available currently and the number that I came up with was I think approximately 88. So there are definitely measures that have been introduced into the mental health assistant since its existence and then removed and will likely be additional measures that will get added as well.

Moderator: Okay here are three related questions. Are these measures linked to some clinical diagnosis and possible treatment? How about the lab results and demographic information, is there a way to select cohort and examine their progression longitudinally.

Carol Malta: Yes definitely; so you know again if you look in that main survey administration table you have your patient identifier, your patient SID and that you can then take and link to all of the other domains in the CDW; so you can link it up to your outpatient diagnoses for example. You can link it up to your labs, you know whatever else that you might be interested in. And in terms of then trying to track over time it’s definitely doable. I mean it does – it will bring up a lot of questions because you know again these measures are all collected in clinical care so they’re not going to care about your timing, right, it’s kind of the timing that just happened so you can have an audit C for example and you’re trying to match it to diagnoses but maybe the time frame is off so that the audit C took place way before a certain diagnosis or whatever it might be. It just – they’ll be data questions that come out or like with labs, you know say you want to look at you know you’re in talk screens for example and you want to see how those results compare with the audit C or with brief addiction monitor. Well the timing of those talk screens is really going to be dependent on when that patient happened to come in for clinical care and all sorts of things that they may not line up the way you necessarily want them to line up. But the possibilities link it all together is certainly there.

Moderator: Is this data being used to assess compliance with recommended screening?

Dr. Hawkins: That’s a really good question. You know it’s monitored – things have changed quite a bit. Several years ago all the screens were associated with performance measures at the sites and there’s quite a bit of – they were scrutinized and at one point in time they were actually tied to serve the compensation packages for like executive directors of facilities. But at this point there’s – I mean the thing to say is they are monitored.

Moderator: Okay, do you get any data from the military health system, clinical data warehouse. Many of these elements are collected by the mental health system on members before they are enrolled in the VHA?

Carol Malta: We haven’t used that data. Not to say that in the future it may be something that we look into though, we can’t speak to it at this point.

Moderator: Okay are you aware and do you use the OEF, OIF surveillance registries created by VHA and managed by the Intellica Corporation; that system has a mental health registry among others such as TBI, poly trauma, women’s health and infectious disease?

Dr. Hawkins: You know I’m not terribly familiar. We certainly used the OEF, OIF registry but that’s just whether or not you participated in that – the wars in Afghanistan and Iraq. But I’m – I guess I’m not familiar with the data registry that’s being mentioned there. I don’t know if Carol –

Carol Malta: No I haven’t; I have not used that in the past.

Moderator: We have a question how do you access a CDW as a representative of \_\_\_\_\_ [00:57:52] post that information for you; so – I’ll do that shortly. Let’s see are there standard screening mental health instruments used in the VA that are not available in CDW, if so what are the common ones that could be obtained via chart abstraction?

Carol Malta: As far as I know everything is \_\_\_\_\_ [00:58:24] some sort of local initiative, anything national should be in a CDW. If it hasn’t made it to the mental health assistant yet it’s typically collected via health factors. So like what’s happened with the BAM is \_\_\_\_\_ [00:58:39] initially it wasn’t actually in that mental health assistant package; so to get the data you would have to pull it from health factors and then the once the switch happened and it got into the mental health assistant then you would pull it off on the mental health domain. But I’m not aware of anything that’s being used nationally that wouldn’t be collected in one of those two ways.

Moderator: Okay do we have time for one more question? On slide 34 survey choice check is not consistent with the survey answer text. Is that what you’re trying to show in the slide meaning we should not use survey answer text but rather use survey choice text, which is more accurate?

Carol Malta: No what I was trying to show is that for your multiple choice responses you’re going to use your survey choice text, that’s where the answer is and for any other data type integers, dates, free text you’re going to want to use the survey answer text and get it from there. So your – the – where the answer lives, the table that it lives in is going to be dependent on your question response type.

Moderator: Okay we have a question who do we contact to do these queries like polling lists of surveys, local folks, do we contact local folks like data managers or some national point of contact?

Carol Malta: Eric I don’t know if you’ll want to take that one?

Dr. Hawkins: Could you ask that question again?

Moderator: Who do we contact to do queries, like polling lists of surveys, local folks like data managers or national point of contact?

Dr. Hawkins: Thank you, yeah you’ll have to find – I’m not aware of any sort of national contact so you’ll have to find individuals locally that are familiar with these data.

Carol Malta: If it is research then you can get some help from the \_\_\_\_\_ [01:00:45] folks –

Moderator: And I’ll address some additional answers to that question or Sara. We have some information that – and I believe that’s all the questions we have. So thank you Eric and Carol for taking the time to present today’s session. To the audience, if your questions were not answered you can contact the presenters directly. Their emails are on the last slide. You can also contact the Vyrec help desk at \_\_\_\_\_ [01:01:18]. Our next database session is Monday, October 3 at 1:00 p.m. Eastern time; the title is Overview of the Data Information Systems, National Databases and Research Uses. And it will be presented by Maria Suden, Research Health Scientist and Associate Director for Communications at Vyrec at the VA Information Resource Center. We hope you can join us.

Thank you once again for attending these sessions – this session. Molly will post the evaluation shortly, please take a minute to answer those questions. Thank you very much.

Molly: Thanks everybody. I’m going to go ahead and close out the session now; so please take just a moment to fill out the feedback survey that will pop up on your screen, it takes just a few moments but we do look closely at your responses and it helps us to improve the sessions and programs. So thank you once again everyone and this does conclude today’s HSRD cyber seminar presentation. Have a great day.