Cyberseminar Transcript

Date: March 2, 2020

Series: VIReC Database and Methods Seminar

Session: Meet VIReC: The Researcher’s Guide to VA Data

Presenter: Maria Souden, PhD, MSI; Jenifer Stelmack, MSW, BA

*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* [http://www.hsrd.research.va.gov/cyberseminars/catalog-archive.cfm](file:///C:\Users\VHASLCMyersK\AppData\Local\Microsoft\Windows\Temporary%20Internet%20Files\Content.IE5\73RXHSNF\l)

Melissa Joyce: Hello everyone and welcome to Database and Methods a Cyberseminar Series hosted by VIReC, the VA Information Resource Center. Thank you to CIDER for providing technical and promotional support. Database and Methods is one of VIReC’s core Cyberseminar series and it focuses on helping VA researchers’ access and use VA databases. Today’s presentation is titled Meet VIReC: The Researcher’s Guide to VA Data. It’ll be presented by Maria Souden. Maria is Acting Director of VIReC where she leads center operations and strategic development to support effective use of VA data for research and quality improvement. Along with serving as an advocate for researchers needs during the VA’s transition to Cerner. She is currently an affiliated Researcher with the University of Illinois School of Information Sciences. Thank you all for joining us today. I’m going to turn you over to Maria.

Dr. Maria Souden: Thank you, Melissa and thanks Heidi. Hi, oh can you see my screen. Am I on? Ope, there we go. All right. So welcome everybody. Thank you for joining us today for Meet VIReC. This is as Melissa mentioned this is a, the fourth in our series our Database and Methods Cyberseminar Series this year.

So we have, if you have missed the first three sessions we started with an overview of VA data, information systems, databases and their research uses. And then we had a session on requesting approval for access to data, and then one on VA Medicare data. So mostly the Database and Methods Cyberseminar focuses on various sources of data in the VA or particular topics and the data that’s used to address them. And some, their application in the research context. Today we have a little departure in that we’re taking a moment to just talk about VIReC. One of the things we find is that people don’t often know what we do. So that’s our goal for the session today is to orient you to VIReC and our products and services.

So this is just a sneak preview of the upcoming sessions. And you’ll notice as we go through the year we have monthly sessions on Mondays, the first Monday of the month usually, that focus on various aspects of using data whether it’s different data sources or data topics. And again you can get to that from our Database and Methods page on the internet or CIDER’s registration pages and then you can always watch archived sessions.

So we’ve said this already.

So what I’m hoping is that by the end of today’s session that everyone is able to describe VIReC activities and be able to articulate our value to the research and quality improvement communities. That you have an understanding of key product types that we produce and how they’re used in supporting work with VA data. That you have a sense of our web environments and can locate resources online. And that you’ll know where to go with, for additional questions or customer service.

So I’m going to, before we get started I have a few poll questions and we’d just like to kind of get to know you a little bit better, know who I’m talking to. So the first one is, focuses on your role as a data user. What is your role in research or quality improvement? And the choices are investigator, PI or Co-I, data manager, analyst, or programmer, project coordinator, or other. And if you select other we’ll ask you to type the specific title into the chat box.

Heidi: And responses are coming in. I’ll give everyone a few more moments to respond before we close the poll out and go through the results. Just waiting for things to slow down a little bit. Okay, I’m going to close this out. And what we’re seeing is 20% of the audience saying investigator, PI, or Co-I, 29% data manager, analyst, or programmer, 30% project coordinator, and 22% other. And in that other category I have IRB regulatory, research assistant, and telehealth coordinator. Thank you everyone.

Dr. Maria Souden: Great. Thanks everyone. That’s great, that’s kind of our typical spread across those types of positions. And you might find that some information here is more helpful to you depending on your role.

The second question asks you to talk about your experience with VA data. So how many years of experience do you have working with VA data? One year or less, more than one but less than three, at least three but less than seven, at least seven less than ten, or ten years or more.

Heidi: And again we’ll give everyone a few moments to respond before we close the poll question out and go through the results. And it looks like we’ve slowed down so I’m going to close this. And what we’re seeing is 44% of the audience saying one year or less, 19% more than one but less than three years, 24% at least three but less than seven years, 4% at least seven but less than ten years, and 9% ten years or more. Thank you everyone.

Dr. Maria Souden: Great. Welcome everyone and definitely welcome to the people who are new to VA data. Hopefully I don’t think that this will get you totally sorted with understanding the environment and the resources but it will, should give you a pretty leg up.

And then the third question just to help us with this particular session is we’re asking you to tell us how familiar you are with VIReC, with our products and services. So the choices here are frequent user, familiar with and use sometimes, are aware of VIReC but never used our products or services, and not at all familiar.

Heidi: And we’ll give everyone a few more moments to respond. And it looks like we’ve slowed down here so I’m going to close that. And what we’re seeing is 4% of the audience being frequent user, 24% familiar with and use sometimes, 37% aware of but never used, and 36% not at all familiar.

Dr. Maria Souden: Great.

Heidi: Thank you everyone.

Dr. Maria Souden: All right. Also really helpful so, all right I’ll try to get everyone oriented then.

So I’m going to start with talking about, really about what VIReC does and who we are.

So VIReC is a resource center supported by health services research and development in the VA. And our mission is really to advance VA’s capacity to use data for research and quality improvement and to, also to form a liaison or a bridge between research data users and the VA health care community. So we have three streams of activity we undertake to achieve this. So the first is that we work with data, VA data to develop knowledge about the data and methods for working with it and using it. We disseminate that knowledge. So we produce products that contain information about data in products and webpages and then disseminate that information so that it’s handy and ready to use when your, when you’re actually starting a data project or starting to work with your data. And the third category of activity is how we improve the data use environment. So we seek to enhance data access and availability for research use. And we do that by working collaboratively with data stewards or other program offices that have a charge around data as well to address issues that come up and to continue to look for additional data resources and how we can facilitate quicker access to data.

All right so kind of how that comes together in our internal structure. So, and this gives you a sense of kind of who’s leading in different areas of VIReC. So we have our data knowledge and data translation group that is primarily focused on producing knowledge about data. We are in the process of hiring someone to lead in that group. We also have a whole stream of work around data access, availability, and use led by Linda Kok. And then we have some cross-cutting work streams that go really intersect and support these two knowledge areas. So dissemination, web, and knowledge sharing led by Jenifer Stelmack. Customer engagement and portfolio management so this includes customer service, needs assessment, and helping to define our upcoming stream of data knowledge work so that it’s meeting customer needs and that’s Chandrea Culbreath. You might know her from our HelpDesk. And then we have our strategic operations work which is really the work done by Melissa Joyce to help keep VIReC running and operating smoothly. So we also have, so all of this and I’m just going to maybe go back one here. So really all of this is supporting the entire data environment for research and operational data users. We also have several special projects that cut across these areas. so again these kind of leverage all of these areas of expertise but are directed in specific ways. So the, our VA/CMS data for research project where we make centers for Medicare and Medicaid data available for Veteran population and Kristin de Groot leads that. Our, the VA REDCap project which is an application for data collection and management that’s run inside the VA and that’s led by Bonnie Paris. We have a special, a stream of activity directed at the needs of QUERI researchers and partnered research. And then this is one that’s kind of within that really that third or liaison category but it’s one we like to call attention to because it’s been very active lately and that’s our EHRM and research project and that’s one that I’m leading here.

So our, this is probably a good slide for reference. It’s the VIReC team. It gives you a sense of what are the key knowledge and functional areas of all of our staff.

And now I’m going to talk a little bit about products and services that we offer.

So I’d like to start, do that by starting out talking about customer service. And Chandrea talks about this really well so hopefully I’ll do it justice. But for us at VIReC customer service is a thread that runs through all of our work efforts. So you can see sort of throughout this whole list of different activities that we’re doing for the community we ground that in an ethos of customer service and being responsive to customer needs.

And what that looks like in terms of our product and service strategy is this wheel or this continual cycle of developing and improving products. So it starts with listening to customers and to our stakeholders within ORD and HSR&D to understand what kinds of priorities need to be supported using data. Then creating documentation and knowledge. So this is where we’re developing products that talk about the data. And then disseminating that knowledge. So we promote that information through VIReC’s websites, intranet, and internet. The VHA Data Portal which we partner on. And then in national conferences, one-on-one through our HelpDesk, and on the HSR&D Listserv. So those are all ways that we get information out and they feed back into how we’re listening. And so then we use all of those, the input from all of those sources to advocate for research use. So as we’re communicating to the community and listening we’re identifying barriers and trying to reduce those. Identifying data needs and trying to advocate for those in the environment. And really increase our knowledge source to address the data that are being most used at any given time.

So I’m going to talk a little bit more about the data documentation that we produce. So we produce a variety of documentation about data. And this can take a number of different forms. So these are some of our most common data documentation or data knowledge products. Our, we have, we produce Factbooks for the Corporate Data Warehouse so these can give a little bit more of that clinical context and some sense of the structure and content of CDW domains. We have a series called The Researcher’s Notebook. Which addresses methodological issues and focuses on how to use VA data to achieve certain kinds of measurements, define cohorts, or look up certain constructs in the data. And then we have products that fall into this sort of data summary realm which is, are really descriptive products that talk about things like, like this, the one that you see here which is describing the domains of the Corporate Data Warehouse. We also describe tables that are available in the domains and produce products that provide counts or detail about contents or historical detail about various data sources.

So this is a little bit of a quick guide to our data knowledge product portfolio. Again this is another one, maybe it’s a good reference after the presentation. It kind of talks about the differences between our different types of data, what they’re trying to get at, which products fall into which categories. And as you can see we continue to add to these products. And this is this, this chart shows the number that are hosted on our website at the moment, in each content area or each genre.

All right and that brings us to our next poll. So this is about VIReC products and I know not that many of you are using our products yet, so but this is for those users. Which of the following VIReC products do you use the most? And the choices are VIReC Factbook, VIReC Researcher’s Notebook, VIReC summary documentation, or other product and again specify that in the chat box. If you haven’t used any VIReC products feel free to select that answer.

Heidi: And responses are coming in. We’ll give everyone a few more moments to respond before we close the poll out and go through the results. Looks like we’re slowing down here. So I’m going to close this. And what we’re seeing is 18% of the audience saying VIReC Factbooks, 8% VIReC Researcher’s Notebook, 8% VIReC summary documentation, 3% other and the response that I got there was VIReC website, and 76% I have not used any VIReC products. Thank you everyone.

Dr. Maria Souden: All right. Well I know now that a hundred percent of you can say that you’ve been to a VIReC Cyberseminar. So I’m sure my numbers will be up the next time I do this.

So I’m going to move on to talking about our knowledge dissemination. So we have a number of venues where we disseminate knowledge digitally. We have our intranet site where most of our products hang, we have the VHA Data Portal which is a site we, that we create collaboratively with National Data Systems, VINCI, and HERC. And those are both available on the, within the VA firewall on the intranet. We have our Listserv which allows for conversations about VA data. And our Data Issues Brief which is our, an e-Brief that comes out usually every two months or so that provides as you can see from the table of contents here, kind of links to news about upcoming events, new products, changes in the data. We try to kind of reach broad and share with you updates from all around the data environment. And then we have our Cyberseminar series. So we have this Database and Methods Cyberseminar Series, but we also have a series we do for, on behalf of QUERI and Partnered Research, information and data systems use and Partnered Research, as well as Clinical Informatics Series as well.

So our digital dissemination environment is composed of kind of like this suite of different venues that really allow us to create a positive user experience and connect data users to information that they need.

And then this reference sheet kind of again, kind of walks through the differences between the different sites. Kind of the public-facing internet is really kind of our brochure online so that anyone can get to. So we do have our, all of our Cyberseminar information linked there as well as linked to peer review publications. Our intranet site is really our main platform for disseminating information. So if you’re using data that would be a site you’d want to circle, star, bookmark in your browser. That’s really where most of our products hang off of the intranet. The VHA Data Portal also is another good one to bookmark if you’re using data. It describes data sources and details the processes for accessing those data sources. And then the Data Issues Brief that comes out monthly and as well as the HSRData Listserv for knowledge sharing.

And the HSRData Listserv is an opt-in Listserv that provides question and answering services for data users. So we have about 1,400 participants. It’s not just people who are actually end-users of data but also data creators. So kind of what we like to think of as the VA data ecosystem. So there are PIs, analysts, data managers, fellows, trainees, as well as data stewards, experts in, various data experts from the resource center. Data architects who are all participating in the Listserv and really answering questions in a very collaborative and iterative way. One of the great wins that we have with the Listserv is it really allows us to articulate what we know as a community and particularly to illicit what we know tacitly about the data. So things that might not be documented in a particularly straightforward way or maybe at a level of detail. Or with a level of historical specificity over time that, where that documentation couldn’t represent every detail of the question. So, but the story can be built by Listserv members responding to a question. And that’s what we see in HSRData. There is an archive where, that you see here, where you can look up previous hosts to the Listserv. It comes right to your email and we are working on developing a more, sort of a more robust environment that can really promote knowledge sharing and better organized information within the archive.

And then we have our VIReC HelpDesk. It’s another main way we disseminate but really more on the individual level. And the HelpDesk is a place where you can call or write with a specific request. If you’ve not found the answer yet on our website or on another website you can write into the HelpDesk. We will triage your answer to an expert in that area whether it’s a particular area of data or other kind of technical expertise. Or we will try to connect you to someone who knows the answer if we don’t know the answer. So sometimes the questions are best answered by another person in the ecosystem. And we maintain a lot of liaison relationships so that we can refer people out and help you get the answers you need.

We also do outreach to raise, keep raising awareness of data resources. So we have what we call our Traveling HelpDesk. You’ll see it at the AcademyHealth Annual Research Meeting, at the HSR&D National Meeting in the years we have one. And at the AMIA Annual Symposium. So if you’re at one of those conferences be sure to stop by and get to meet us in person, pick up some of our literature. It’s a great place to meet up with VA people at those conferences as well.

All right so now I’m going to interrupt your happy listening for a poll. So this poll is about VIReC’s services and our dissemination sites and we’re asking which of the following services or sites do you use the most? And the choices are VIReC websites, the VHA Data Portal, the VIReC HelpDesk, and the HSRData Listserv. And then there’s also an option for I have not used any of the above.

Heidi: And responses are coming in. We’ll give everyone a few more moments to respond before we go through, before I close it out and we go through the responses. Looks like we’re slowing down so I’m going to close this. And what we’re seeing is 23% of the audience saying VIReC websites, 29% VHA Data Portal, 4% VIReC HelpDesk, 24% HSR&D Listserv, and 49% I have not used any of the above. Thank you everyone.

Dr. Maria Souden: Great. Okay, great. Good to know. And then the next poll and I think this is my last poll for the day is, which resources are or do you think would be of most value to you in your role? These would be data so, and the choices are information on data sources and accessing them, information on recent changes in the data environment, information on special topics and if you’ve got topics in mind we welcome you to specify those in the chat box, training in skills like SQL and SAS, or training in project management.

[silence 26:20-26:48]

Heidi: I’m sorry I was muted there.

Dr. Maria Souden: I was going to ask.

Heidi: Thank you. The options here 81% of the audience saying information on data sources and accessing them, 31% information on recent changes in the data environment, 7% information on specialty topics, and we didn’t, special topics PIT and Cerner, 51% SQL or SAS training, and 36% project management training. Thank you everyone.

Dr. Maria Souden: All right. Great, thank you. We’ve never asked that question before in this seminar so it’s really definitely interesting to see what kinds of resources you all are looking for. And what kinds of ways we can continue to add value. So after this little tromp through our dissemination in health venues, I’m going to talk a little bit about how we bring that together in our role as liaisons.

So we’ve been monitoring the needs, we’ve been documenting data and responding to needs by creating knowledge. We’re disseminating that knowledge and then we’re using feedback in all of those phases for our advocacy or liaison efforts in the community and on behalf of researchers.

And I’m going to talk a little bit about some of those liaison efforts before I dig a little more deeper into some like use cases for our products. So one of the things in this goal area that we try to do is we really are, a lot of our work centers around improving access to data and availability of data. So if we know there’s data that people really need or there’s an opportunity to make more data available for research use we undertake initiatives to try to do that. Or if we’re, we try to improve the process of getting access to data so that the time, that time from like putting in the request, to actually got my data and doing my work is shortened over time. And one of the main ways we do that is by participating in the Data Access Request Tracker or DART team. So on that team we work with VINCI and we work with National Data Systems to do a few different things in service of customers. So we try to increase the number of data sources that are available through the DART application process. So you can see these are some of the new data sources we added last year. We work with data stewards and program offices that generate data trying to help them kind of problem-solve about how to make their data available and how to provide the best service to data users. And then we continue to take on improvements in the DART system. So the team works at improving usability and improving the forms and the submission processes. So that we’re creating a better experience not only for the data requestor but also for the data stewards who need to approve all those requests on the back end.

Another way that we contribute to availability of data in the environment is by serving as the steward for VA/CMS data. So this is the only data that we actually have, that we actually put our hands on. The rest of the time we are helping to understand data that exists in other places in the organization. But we do have an agreement with the centers for Medicare and Medicaid and the [unintelligible 30:39] for enrollment, policy and planning maybe, around, that we are taking the data from centers for Medicare and Medicaid and make that available for research use. So we have kind of three pillars of activity in this project. So we provide data, we provide education and assistance in using that data, and then we provided governance over the data. So it’s a little fuller than our role with some of the other data sources which allows us really to not only provide the service and be a good steward of the data for VA but also it helps us to inform your experience in some of these other areas, with other forms of data. So we maintain a repository of data, Medicare and Medicaid, U.S. renal services data. We provision data that’s requested. We also contribute to the data resources of the organization by, we’re undergoing a conversion of the CMS data to the OMOP common data model, working with VINCI on that. We provide education and assistance using these data. So we produce some products that help people understand what’s in the data and how best to use it. We do Cyberseminars and we do, we have a HelpDesk for help as well as doing pre-request consultations. So if you’re using VA/CMS data we’ll ask you to have a call with our experts so that we can make sure you’re getting what you need to meet your research aims. We also run the Axis request process for data. Take care of policy governing the appropriate use of the data. These data are considered pretty sensitive in the VA environment so there’s a lot, there’s some additional constraints around their use. And then we continue to maintain oversight as long as projects are still holding data.

This kind of gives you just a brief snapshot of the data that we make available through the VA/CMS project. It’s about five terabytes of data, about 2,600 SAS datasets. And covering the, from, since 1997 so about 20, well just over 20 years of data. And that’s from Medicare, Medicaid, the renal services. It includes everything from claims to assessment data to survey data. As you can see we’ve added some new data resources in the last year including quarterly data, as well as just newly in Medicare and Medicare Advantage encounter data. And we’re still looking at getting, increasing our data resources. So VIReC does have a section on our website all about the VA/CMS data for research projects which provides information about the data that are available and the process for getting it.

So these are some of the products that we produce. Talking about VA/CMS data. You’ll see us presenting at conferences, usually we’re trying to do comparisons with VA, general VA health care use versus use of Medicare and Medicaid by Veterans. Sometimes we’re comparing populations. We produce data snapshots that allow like kind of a peek inside the contents of the data as well as a lot of technical documentation about what’s available and the data contents.

Another way in which we contribute to the data use environment is through hosting VA REDCap. So REDCap is Research Electronic Data Capture software and this is a web-based application that’s hosted by VINCI and that VIReC makes available for research data collection and data management. So it’s free to use, it’s secure. You can learn more about it from the VIReC homepage. But it is a service we offer to help collect and manage primary data.

The last thing I’m going to talk about in this area of liaison is the work that we’ve been doing for the Office of Electronic Health Record Modernization. So we’ve been, for about the last oh maybe two years I’ve been, along with Jim Breeling who’s the Director of Bioinformatics at the Office of Research and Development co-leading the research workgroup for the Office of EHRM. And this is the group that specifies functional requirements for the Cerner system that are needed by research. So this is just kind of a, just a snapshot of the most active people on our workgroup. You can see that they are, they represent across the services. This is an ORD-wide initiative. And it’s really to make sure that as we transition to a new Electronic Health Record system the commercially available system offered by Cerner that we are still able to maintain the kind of research access we need and then the data that comes out of the record. Kind of no small feat so we have focused in the last couple of years on really developing an ecosystem of activity around research and the EHR. So you see our research workgroup at the top and then we own those couple of layers underneath that. And really look at through these efforts we try to reach down to the field. So we kind of have this tiny point at the top like where we’re actually able to influence decisions that are made in the implementation of the health record, but we have lots of support underneath. So we go abroad to get the best expertise from the field. To get escalation of needs and issues. And then we reach back down to provide updates, education, and guidance. So that’s really just getting going, it’s kind of been a lot of preliminary work so far but you could expect to hear more and more from us over the next few years and on throughout the rollout.

So some of our specific purpose work efforts that are going on. So this is like those mapping those top three layers of the pyramid. The research workgroup works with the Office of EHR Modernization to represent our functional requirements and needs for the EHR and data coming from the EHR which is basically the equivalent of CDW. The ORD Strategic Initiative for Research-EHR Synergy or OSIRES is a group that I am co-chairing along with Jennifer Sporleder from the, from CSP or the Cooperative Studies Program. And that group is really an ORD-centered group to help continue to provide long-term strategic planning and mitigation for research's effort to maximize the benefit of the EHR and the modernization of it. And then we have in purple these next I think four layers or four initiatives on the purple layer are all subgroups of the research workgroup. So they focus on different aspects. So the CSI-Seattle team looks at risks, gaps, and priorities for research at the initial operating capability sites. Puget Sound is the first research site to go to Cerner and that will happen later on this year. We have a clinical trials team that focuses on configuring Cerner’s power trial solution and facilitating the implementation of that into the Cerner EHR. We have a research data sustainability team which is engaged in activity around making sure that we are getting data out of the system that’s comparable to what we get now like through the Corporate Data Warehouse. And making sure that that data comes from Millennium, the Cerner EHR back to the VA Central environment and that we understand how to use it along with some of our legacy data. And then we have our cooperative research and development teams. And these focus on specific projects that allow us to collaborate with Cerner to develop and innovate in new areas within the Electronic Health Record system. All right so that was the like, the big broad trump through all of VIReC’s areas of activity. So you can see we’re keeping busy.

The next section, the next kind of last bit I want to wrap up with are talking about what our data resources look like in action. And we’re kind of inspired to have this section because one of the biggest questions that we get from the field is like where do I go to do this. And we do, we live in an environment that’s pretty disparate. So there, the data use environment base is distributed. There are a number of different stakeholders. So we, what we try to do is kind of help map that landscape and I’m going to talk today at just some of the general use cases for VIReC products in the landscape and where you can find those products.

So again this is not a detailed exploration but should give you a taste of the kinds of things you can find on our site and their usefulness in different research activities. So just to review kind of our sites that we, where we host data knowledge or provide data knowledge resources. The VIReC intranet which is only accessible when you’re on the network and that’s, this is where a lot of our resources for working with data hang. It’s organized by sources, by topics, particularly data, our products are organized by the data sources that they represent. And then we have representation of topics. So if you were interested in understanding how you might take on looking at [unintelligible 42:07] care in the community data, you might go to our community care data page. Or if you’re trying to understand race and ethnicity, you might look at the data topics page for it. I think I’m scooping myself for a little further in the presentation. But the second intranet resource we have is the VHA Data Portal which I mentioned before. So if you, this has information about data sources, data access. It’s aimed at both research and operational uses, users of data. And really addresses some of those technical aspects of selecting a data source and getting to your data. And then we have the VIReC internet site which again is our public-facing site where you can find information about, you can find links to publications as well as to our Cyberseminar schedules and archives.

So the use case number one I want to talk about is just getting the lay of the land. So this is the like huh what’s out there, what kind of data does VA have. And for this the Portal is an excellent starting place. It’s a good place to get an overview of VA data sources. Also might be helpful used alongside of the, our first Cyberseminar in this Database and Methods Series, so the first one we do every October is the overview of VA data. So that kind of paired with the information that’s indexed here on the VHA Data Portal gives you an idea of all of the data sources we have. If you click on any of these links you get taken to descriptions of that source. As well as more specific information about how to access those data.

Use case number two, you want to know about contents of a particular data source. So I’m picking on CDW here because it’s really our biggest data resource. It’s probably what most people use when they’re using secondary data that’s been collected in health care operations. This is our CDW documentation page. So this is where you’ll find links to all of the documentation we produce about CDW. As well as links to documentation or knowledge that’s produced by other players in the data environment. So like the VINCI site or the CDW/BISL SharePoint site which has a lot of metadata as well as tools that provide you with counts of, of things like missing data as well as links to additional training. So this is a good starting page kind of resource to get to a lot of our CDW documentation products.

Whoops, oh! So some of that documentation I talked about these a little bit before but this is a closer look. So this is the, this is a Factbook. You can see that it walks through kind of the clinical context of how the data got into this domain, where do they come from in the clinic, how are they transformed, how are they labeled and organized. What are the primary and the foreign keys you’re going to use here? And then some, provides some example SQL code. Just to get a sense of how the elements in the domain work together.

So the next use case is technical aspects of working with data. So this is like maybe I’m a little further into my project and I’m trying to do, opes, oh yeah okay. This is actually beginning, never mind. I thought I was on a different slide. So this is the, this is to really understand kind of the mechanics of working with CDW. So we have a suite of CDW Cyberseminars that are archived and made available in this little library on our website from that documentation page. It really is our getting started library. So these products really step you through not only understanding the Corporate Data Warehouse but also understanding how to use SQL along with the Corporate Data Warehouse structure. So I think it provides that kind of integration of like the technical skills with the knowledge of the data contents. It’s so helpful for getting started in this environment and with these data.

And then there’s resources around the conceptual aspects of working with data. So these are resources that help you focus on specific topics, specific measurements, and specific methods. So these are those data topics pages that I was previewing before. So you can see some of, just some of the topics that we have pages about. And usually they will talk about various considerations for using these kind of data. Sources where you might find data in these areas. Our Databases and Methods Cyberseminar Series are good introduction for may topics. So these kind of run the gamut from pharmacy to race and ethnicity data to comorbidity Cyberseminars.

And then when you’re ready to kind of get in there and really work in the weeds with the data we have our Researcher’s Notebook Series which drills down in kind of one very specific question like the example you see on the screen, how to identify bed sections for my cohort. And then steps you through the process of achieving this, putting together the tables that you need to, understanding their logic and how they go together, what views you actually need and then how to write the queries and some sample code.

The next use case is another one for the VHA Data Portal. So this is kind of maybe goes, maybe in the process should have gone like right there after finding out about your sources or getting the lay of the land. But this is how to get to data. So data in the VA you need to apply to use it and you need to meet certain criteria. Linda’s archived session in this series, the second session, the November session focuses on some of the rules around accessing data. And our VHA Data Portal site details the access processes and criteria for each data source that we talk about here.

All right so likely you still have questions. Please I hope you go and check out our websites. We’re going to, we have a few minutes to take some questions from the audience but if there’s still lingering things we invite you to reach out to our HelpDesk. You’re always welcome to send an email to me although Chandrea might be a little more helpful to you. So this is our contact information.

And I just want to mention that on the last slide in the packet here is a list of quick links. So this is another good takeaway resource. Just provides links to some of the resources we talked about today. The first one is actually a PDF that has a lot more links in it. So that’s a whole collection of quick links to resources for using VA data and that is published on our intranet site as well. So I’m going to stop here and take questions if people have them.

Melissa Joyce: Hi Maria. Thank you so much for your presentation today. Quite a few questions have already come in for you. So I’ll start with the most popular question. I think I’ve gotten it from three different people already and that’s the question about the relationship between VIReC and VINCI. Can you speak to the relationship between VIReC and VINCI? And can you explain the difference between the two teams with regards to data sources and accessing them. And then just how do they overlap comparing and contrasting them. So all regarding the VINCI/VIReC relationship.

Dr. Maria Souden: All right. Well that’s a good meaty question, all right and we do get that a lot. So VINCI, like VIReC is funded by Health Services Research and Development, HSR&D, to provide resource center services to the data using community. VINCI by the way stands for the VA Informatics and Computing Infrastructure, I want to say. So VINCI’s origins are really in providing an environment and providing an infrastructure for using the data. So kind of where you see the roles coming in is when you actually, so as you saw in the presentation we collaborate with VINCI on a number of things. So a lot of resources that support using data we collaborate on. So like the DART system for researchers. And the, in the course of finding information or figuring out what data you need to request for your project you could get that help from VIReC. You could probably also get it from VINCI. But you just might get it at different points. So usually where the VINCI concierge might step in is once you have your approved DART request and you’re ready to have some data provisioned to your workspace that’s where you would get that one-on-one help in kind of figuring out how to operationalize your cohort in the data and your outcomes of interest. The, our focus is really on understanding the data. It’s on data knowledge, data education. Their focus is on providing an environment to use the data and the infrastructure for that. They provide the secure workspaces. They do a lot of hosting of data. And then of course as you can imagine there’s overlap between these two things. So the actual processes around getting access to data, the knowledge that goes around the use and requesting of data. Those are present really in both of our organizations and they’re also places where we try to collaborate and synchronize as much as possible. So I think I maybe answered the question. I’m trying to think if I got to everything. But you’re welcome to follow-up with me if you want more detail on that.

Melissa Joyce: Thanks, Maria. The next question that come in is regarding the Cerner transition. How will the gradual transition of Cerner affect national EHR data use? i.e. Puget Sound is one of the sites involved in a research project so they will transition to Cerner long before we do here on the east coast.

Dr. Maria Souden: Right, so that’s definitely correct and that’s, I’m spending a lot of my, a lot of my energy, a lot of my time working along with that workgroup that you saw up there to try to make this transition smooth for research. And certainly one of our goals is to make sure that we continue to have access to a complete national set of data. So as I think the questionnaire correctly hypothesized like once, once that data is being collected through the Cerner Millennium EHR, the VistA system is no longer and CPRS is no longer up and running at that site. So when Seattle goes then all of the data sources that have national-level data do not have Seattle data anymore. That’s collected after a certain date. So after December you won’t see encounters from the Seattle sites of care. And so we are working on a, sort of a bridging mechanism for that and we’re working with a lot of other people so this is a pretty big ecosystem in and of itself. We’re working with the Office of EHRM and the technical organization there. So the data management and migration team is led by Jack Bates the Director of BISL, the Business Intelligence Service Line. That, BISLs the architect of the CDW data. And so they are architecting the syndication of data back out of the Cerner environment into a centralized analytic environment where it can be used along with CDW data. So it’s likely that initially we’ll have some Millennium formatted data and we’ll need to really learn how to discern where the two data models are alike and different. Eventually I think the BISL team will be working on developing a converged model of CDW so that it can have both data represented in it.

Melissa Joyce: Thanks, Maria. The next few questions are related to finding help and finding resources, looking for specific direction. So the first one is, if a researcher wants to do a chart review using VA data where do I direct them?

Dr. Maria Souden: So on the, the Data Portal there is a section called tools. And we do describe there some of the tools that are available for working with data including, so previously the CAPRI tool you could use for reviewing records at a single site. And then we, for a long time used VistAWeb to look across sites and now that’s being phased out or has been phased out and we were using the Joint Legacy Viewer, JLV. But that’s, those are the main tools for chart review. There’s also a chart review tool that’s available through VINCI as well that does some more NLP intensive kinds of reviewing and extraction activities. But that information about those tools, how to get access to them is available through the VHA Data Portal. We usually do, try to do a Cyberseminar every year, I think we have one coming up in July, as well as we’ve got some archived sessions as well that kind of provide an overview of the chart reviewer tools.

Melissa Joyce: Great. May be time for one quick final question as we’re just about to the top of the hour. But it’s, are there any ways to find or even hire a data analyst who is familiar with VIReC or CDW?

Dr. Maria Souden: Yeah, that’s a really, I mean that’s also, a lot of these are very common questions that we get. There isn’t a, like a formal institutional pathway for doing that. So I think people make, do make connections through the HSRData Listserv. I know for a while VINCI was piloting kind of a matchmaking database that would allow you to locate data analytic expertise. But typically I think the way that these arrangements usually happen is that, is through networking and finding people who have experience in the topics that you need or have skills working with certain kinds of data. And then with arranging to bring them on as part of your project.

Melissa Joyce: Great. Okay, thank you Maria. As we’re just about at the top of the hour I want to close the question and answer session for today. But encourage you all if you still have questions, please as Maria said contact the VIReC HelpDesk or Maria directly. I think we got through most of the questions today but I think there’s a few remaining so we’ll try to reach out to those of you who’s questions we weren’t able to answer today. But in closing, Maria thank you for taking the time to present today’s session. Please tune in for our next session in VIReC’s Database and Methods Cyberseminar Series on Monday, April 6th, at 1 p.m. Eastern. Wyndy Wiitala and her team will be here to present, Weight, Weight, Do Tell Me: Comparing algorithms to process and clean CDW weight data. We hope you’ll join us.

[ END OF AUDIO ]