Cyberseminar Transcript

Date: May 10, 2018

Series: VA Informatics and Computing Infrastructure

Session: VINCI Data and Data Services

Presenter: Kevin Malohi

*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)

Rob: And as it’s just the top of the hour, I'd like to introduce our speaker today, Kevin Malohi, who is the VINCI data services manager, and his team. They’ll be presenting today. Kevin, can I turn things over to you?

Kevin Malohi: Go ahead, Rob. [Unintelligible 00:23] Can you see the screen?

Rob: Yes sir, it looks good.

Kevin Malohi: All right. Thanks for the introduction and welcome everyone to another VINCI training series today. Please keep in mind that the focus of this presentation will be on VINCI data and what’s available within VINCI. There are more than 62 data domains that’s available, and so we won’t have very much time to go into each domain in detail. However, we do have references and hyperlinks that you guys can refer to at the end of these slides. Okay. What happened?

Rob: If you can’t move your slides, Kevin, you should probably just click on the slide, and then you’d be able to use the keyboard. There you go.

Kevin Malohi: All right. Presenting today will be the data services team starting with me, followed by Steve Oostema, Terry Crowe, Elise Gatsby, Lacey Lewis, James Potter, Brian Robison, and Tori Anglin. In addition to some of the services we provide our VINCI customers, which we’ll cover later in this presentation, we also work with a number of VA programs to provide a myriad of other data support as well. And some of these programs include the Million Veteran Program, the Cooperative Studies Program, and a informatics group called MAVERIC, which is based out of Boston.

So looking at the agenda for today, as I mentioned earlier, the bulk of this presentation will cover VINCI data. Now there’s more than 1,000 tables, somewhere around the neighborhood of 25,000-plus data columns, well over 72 terabytes worth of data, so it’s going to be virtually impossible to cover those in detail today, so we’re just going to provide a very high-level overview. In addition, we’ll talk about some of the services and limitations that we provide, and then most importantly provide you guys with tools and resources that you guys need to be successful. And then time permitting, we should have a Q&A session afterwards.

All right, so here’s the first of two polls. We just want to get to know our audience better so we can improve these Cyberseminars in the future. Rob?

Rob: I’ve launched the poll, and answers are streaming in. We have about 20% voted, so I’m going to give people a little bit more time to go ahead and answer. But the question, what is your role in VA? Answer number one, researcher investigator/PI; answer number two, data manager, analyst, or programmer; answer number three, project coordinator; answer number four, clinical or operations staff; and finally other, and you can use the questions pane to describe what other means. We have about 80% voted, Kevin, so I'm going to go ahead and close the poll and share out the results. And what we have is that 28% of your audience chose answer number one, research investigator/PI; 32% chose data manager, analyst, or programmer; 9% chose project coordinator; 21% clinical or operations staff; and 9% answered other; and 1% said that they are VINCI concierge. I think you probably know who that person is. Would you like me to continue on with the second poll?

Kevin Malohi: Yes, go ahead.

Rob: So that poll is up. And to the question how many years of experience do you have working with VINCI data, answers are streaming in. It’s looking like the majority have less than one year. Answers are flattened out at about 85%, so I'm going to close the poll and share out the results to the audience, but I’ll tell you that 66 say that they have less than, 66% say that they have less than one year experience, 25% say they have one year to less than three, 5% chose three years to less than five, 4% five years to less than seven, and zero people chose seven years or more. So we’re back to your slides now.

Kevin Malohi: All right, thanks Rob. Seems like we do have a good mix of users as well as people with experience, and the bulk of our listeners are fairly inexperienced with VINCI. Thank you.

All right, as far as the learning objectives we want everybody to take away, one is a better understanding of the data that’s currently available on VINCI and then knowing what services that we do provide, and then where to find the tools and resources to be successful. All right, so for the sake of this presentation, the data has been divided into four primary categories starting with production data, raw data, and SAS-based datasets, and we also have operational projects as well. But for the sake of this presentation, we’ll just focus on the top three.

All right, moving into the production data, we’ll have Steve present this portion.

Steve Oostema: Thank you Kevin. CDW production data, CDW stands for Corporate Data Warehouse, and the vast majority of this data does come from CPRS. There are basically two types of tables found within CDW production data, dimension tables and fact tables. You can connect back and forth between these tables and join them together to get your cohorts using unique identifiers such as station identifiers. Those will have the marker SID within the column name. Examples of those are patient SIDs and ICD SIDs. There is also one unique identifier per patient across the entire VA, and that is known as a patient ICN. It should be noted that production data is updated nightly. Next slide.

The first thing to look at is the dimension tables or NDim tables. Dimension tables do not contain any patient data. They contain more information about descriptions of different things such as ICD diagnosis, numbers and description names, as well as drug classes. NDim tables is a schema for any dimension table that is not sourced directly from a VistA system. Next slide.

Fact tables in general contain patient information. They can be provisioned to research studies by request, and most fact table data does come from CPRS.

The first domain we’re going to look at today is the allergy domain. This contains information about causes for hypersensitive reactions. Includes medications, foods, chemicals, and noxious stimuli, and there are six main allergy data tables and one dim table associated with the allergy domain.

The appointment domain basically contains information about appointments. They are future upcoming appointments as well as past appointments. You can also find information on the staff member who is involved with the appointment. You can find information on who the patient saw during that appointment.

The consult domain contains information about requests for services. You can find information on specialty care and procedures there.

The CPRS orders domain contains information about any order that a doctor may have given for that patient. This could include labs, radiology, and medications, although it should be pointed out as far as medications are concerned that this will not include any results for the labs or any actual medications that were listed that the patient actually picked up.

The dental domain is actually one of the largest domains within CDW, but it is not used very often. You can find basic hygiene, exams, orthodontics, periodontics, endodontics. Any dental surgeries that are done can all be found within these tables. There are 29 data or fact tables and 17 dim tables associated with the dental domain.

The EDIS domain, which is the Emergency Department Integration Software domain, is where you can go to find information as patients were tracked as they went through the emergency department.

The health factors domain includes information on over 300 different types of health factors that patients may have assigned to them. It includes items like tobacco use, alcohol use, and drug use. Not every VA station uses the health factors domain, and so as such, it’s probably not as full as what you would think it could be. A lot of health factors are often used with clinical reminders. The data does go back to fiscal year 2000, and there is one data table and one dimension table associated with the health factors.

The immunizations domain contains records of any patients who have gotten immunizations or vaccinations through the VA system. There is one data and fact table, and there are no dim tables associated with this domain. There’s a [unintelligible 11:50] called VisitSID, and basically what that is is the unique identifier for that patient at that station, and that can be linked back to outpatient visit records. Outpatient records will be talked about later in this session.

The inpatient domain has a huge selection of variables that you can look at. Basically it pertains to any inpatient status that a particular Veteran may have. Includes admission dates and discharge dates, ICD or CPT procedures, any surgical procedures. You can also find the ward location where the patient was located at. You can also find out whether or not the patient had any major diagnoses, and there are primary, secondary, tertiary. There are several different diagnoses listed for each inpatient stay.

Terry Crowe will now cover the next slides.

Terry Crowe: Terrific. Thank you, Steve. All right, so the next data domain is lab chemistry. This dataset contains laboratory results for patients. Location information about the physical institution, physical laboratory, session area, and lab section is found here. Details about the lab tests, the test itself, whether the test qualifies as workload, whether the test is part of the panel, as well as the specimen collected for the test, the SNOMED code, LOINC code, and national VA lab codes are included in this dataset.

The MCA, formerly known as DSS, this is the managerial cost accounting dataset. We find information to help us to figure out budget impacts as well as total costs, costs of care for a patient or groups of patients, and this data can be linked to inpat, outpat, and pharmacy data.

Mental health assessment dataset. Here we have information related to all mental health assessments and assessment instruments. User can find information related to the type of assessment, instrument, the purpose of the instrument, the name of the instrument, the questions included in the instrument, and the answer options for each question. As well, you’ll find the point systems that go with each of the answers, the date that the test was administered, and the overall scores on the instrument. Next slide.

Microbiology. Microbiology is a dataset where you’ll find information on bacteria, viruses, and parasites. This contains information such as specimen dates, locations, requesting physicians, and there is a lot of information in this dataset in CDW.

Next we have non-VA meds. The non-VA meds dataset contains information about the medications that were prescribed outside of the VA, and there is data on the details of the prescription such as medication strength, frequency of route administration. There’s also dates of use and when the medication was entered into the system.

OMOP. This is the Observational Medical Outcomes Partnership framework. This is a standardized format that helps perform systematic analysis of disparate databases. This dataset is, the OMOP framework has been around for a number of years, and it allows to have a strict set of conventions on naming, content, and the use of standard vocabularies. The data is mapped to a controlled vocabulary in order to be computable. So this groups some of the data that we’ve got in CDW in a different way than the CDW framework. Makes it helpful for a lot of the studies and gathering information in particular ways.

Next slide is outpatient. This is also one of the largest datasets in CDW. Contains details of outpatient encounters, outpatient visits, and inpatient encounters. Data are collected and assembled on a variety of aspects, these episodes of care including the logistics of the appointment, problems reported by the patient, the exams and tests executed, and diagnoses rendered by the providers, and also the procedures performed during the visit timeframe.

Patient associated. This is a table that is not used very much. It’s actually got one purpose, and that is to have information regarding the radium exposure for patients. It’s a very specific table, and it gets used very little.

Patient enrollment is information on the patients as they’re enrolled in the VA medical care system. As soon as the patient joins the system, they get a record in this database, which gives all the information on the patient for future use in other tables upon their enrollment. Pretty simple.

Now we will hear from Elise Gatsby.

Elise Gatsby: Thanks Terry. The first one I'm going to touch on is the PCMM, Primary Care Management Module. It describes the team of providers and their relationship to an assigned patient, things like the provider title, role, the service section, team purpose, and some dates. The primary tables here would be patient providers, team assignment, preceptor relationship, and PCMM contains data only up until the time the station converted to RPCMM, which I'll talk about here. It’s reengineered PCMM and incorporates these patient provider relationships from a third-party system as well as the VistA domain. And as of last I checked, 130 stations have transitioned, so all of them have made it into RPCMM. You can check the transition date in this VistA transition table, and then you can know based on your date which table you might want to look in.

Next is PCOR. It’s from PCORnet. It’s their common data model. They have a lot of documentation on their website, which we will provide at the end of this slide set. Some of the main tables of interest is this HARVEST table, which describes the implementation and the data refreshes, a crosswalk so you can link the PATID back to something the VA uses like a ICN or a SID, a table dedicated to patients enrolled in PCORnet trials, and then a table dedicated to these patient-reported outcomes that PCORnet uses.

The pharmacy BCMA, bar code medication administration, describes medication administration for inpatients. Dates and times of medications ordered, delivered, or administered, details about the medication, any notes on whether it was refused or skipped, and information about PRN drugs. The main tables are the medication log, and then from there you can link to dispensed drug, additive, and solution using the medication log SID.

Pharmacy outpatient is exactly what it sounds like, outpatient pharmacy information. Includes some information about the administrative services of the pharmacy and then information about the patient and the prescription. There is not a distinct field for dose, unfortunately. You’ll have to get that using a different field which includes name and dose. And there is specific information about clozapine if you’re interested in that.

Purchase care, also known as fee, is reimbursement data for non-VA medical care. It includes the vendor, the reason for the visit, dates of service, and then the financial information like the amount billed, the amount paid, and the type of payment. There’s also some clinical information such as diagnosis and discharge/admission dates. There’s quite a few tables in this domain. Some of the main ones are fee authorization, fee service provided, and then a couple inpatient code tables. And with this table, it comes from VistA. There’s also a fee basis claim system that’s separate, which Lacey will actually talk about later.

Radiology was moved into production in July of 2017. Includes information, clinical or administrative, about radiology, exams, orders, and reports. The rad schema is the fact tables, and then SPatientText schema includes the report text including notes and clinical histories. So some of the main tables are radiology exam and then the SPatientText. The file structure is set up to allow for imaging data, but as far as I know, there is none as of yet.

SPatient and patient, this is all patients followed by any medical center or outpatient clinic in the VA, information about the patient themselves, so demographics, sensitive patient information, and contact information. There’s some military service info such as service-connected disabilities or period of service. We don’t have things such as deployment or service location. And the main tables here would be patient or SPatient, address and phone, and then military service episode. If you do ask for this as part of your project, the screened or scrambled SSN and real SSN won’t be provided unless your project specifically approves of those.

SStaff and staff is similar in that it’s information on all VA staff. Primary tables are staff and SStaff, provider type, and VistA permission. And again you won’t have access to real SSNs unless your project specifically is approved for that.

And last I will touch on surgery. This is data extracted from the surgery package. It’s split vertically into three conceptual tables, pre, intra, and post, which are exactly what they sound like. Pre would be surgical scheduling and prep, intra is data collected during the surgical procedure, and then post is the notes, the indications, and anything about implanted prosthetics. Each of these tables does have enough info to provide for any basic query, so each of them has case identifiers, patient identifiers, date and times of the surgery. And then beyond those three, there’s specific tables for associated diagnoses, procedure codes, and CPT codes.

And from there, I'll turn it over to Lacey.

Lacey Lewis: Okay. Thank you, Elise. The next domain that we’ll discuss is women’s health, which is exactly what it sounds like. It’s specific to the care of female Veterans, and there are the associated domain tables. It is a fairly new domain and one that is being highly used in terms of grants and different RFAs that are being submitted.

Next, I’ll discuss vital signs. This has original datasets available on CDW and VINCI and contains vital records. It is one of the largest tables with billions of records. It has different measurements associated with different items such as pulse rate, blood pressure, height, and weight. There is one thing to note with this in that there is no unit, and so most of the units as discussed here are in, for example, weight in pounds, respirations breaths per minute, and the BMI can be calculated based on the height and weight.

TIU is another domain that is now being highly used. It has all of the clinical data that comes in the text and notes from providers and contains all of the clinical notes. It is one of the largest domains in CDW and VINCI in terms of its size because of all of that information that it contains. These are the domain tables that are associated with TIU. Another thing to note is that notes are full-text indexed, making it very easy to search for particular terms that are mentioned within clinical notes.

Travel. It contains patient travel information, so it contains their information of travel to a care facility or that they have been transferred with that travel to another center for care, and the associated domain tables that you see listed here.

Next, we’re going to be discussing raw data, which is separate from the production data that we’ve covered. So the first raw domain that we’ll be discussing is fill claims. This provides billing information for care outside of the VA. It’s not a common domain used amongst researchers, but it is available if you’re wanting more information about those claims.

Next is echocardiogram, and this contains the echocardiogram procedures as you can see the different following domains amongst many others that are provided. And this information also may be found within the TIU domain.

Left ventricular ejection fraction information is a very new domain. One thing to note is that it has been provided from our VINCI team through natural language processing extraction and gleaning of patient notes. It does provide that left ventricular information. It has the date of when that measurement was taken and then also a range that includes the lower and upper values of that dataset.

Okay, FBCS is the fee basis claims system. It’s a nationwide system used to improve claims management, and this is a non-VistA extraction. There are 42 tables under this domain.

Okay, inpatient medications. So this information is in regards to any type of IV given to patients. It contains the type of drug, mixture, dosage. All of this information is included within the domain. Its counterpart is the inpatient medication unit dose. However, if you want more information, you can also request the pharmacy outpatient or BCMA medication domain to obtain any other information beyond the IV medications domain.

Lastly, I'll talk about oncology, which automates tumor registry and supports tumor registrars. This information, specifically the domain 165.5, that’s been a domain that’s been highly used from different investigators and includes information on site of metastasis and SEER summary stage.

I will now turn the time over to Jim who will continue to talk about some of the different raw joint data domains.

James Potter: Thanks Lacey. The first one I'd like to talk about is pulmonary function test, which is one that is not used very often. I've only ever had one customer ever request anything from pulmonary function test. It represents one patient fact table. It comes directly from actual pulmonary tests, and then two lookup tables that essentially are like dimension tables.

Fifty-two, prosthetics, [unintelligible 29:46]. Prosthetics tracks everything that goes on or into a person or a person gets into. There’s actually a vehicle table that tracks real vehicles like trucks, vans, and automobiles. Prosthetics tracks really surprisingly detailed set of items like corneas, [unintelligible 30:13] expanders, dental implants, breast and penile implants, and the items to order are found by HCPCS codes rather than CPT codes or the ICD codes you would normally expect to see. There’s one dimension table, and that’s dim [unintelligible 30:30] prosthesis that contains the orderable items. Next.

Veterans Choice, or the VACAA database. This database was passed as a reaction to the problems we were seeing in the VA back in 2014 where patients were found to be waiting months for a chance to get in to see a specialist. So what they’ve done is they’ve created the Access To Care Legislation in 2014 that lets Veterans choose healthcare providers in their own community, and then there are four rules that limit the choice option for Veterans, which these four rules have been in some cases not very popular with the Vets. The distance is 40 miles. If you live over 40 miles away from a station, you can use a local provider. If you can’t schedule an appointment within 30 days, you can have a local provider. If you have difficult travel conditions like across water, airplane, even walking, horse, whatever, then you can get the Veterans Choice. The distance of 20 miles in Alaska, Hawaii, New Hampshire, or US territory except Puerto Rico, which I'm not really fully understanding of why, but that’s the distances that we have, the distance for those people.

Okay, next, care assessment need, CAN score. This is a score that actually has been used quite a bit. It’s from zero to 99. It’s probability of a patient going into hospital admission or death within a specified timeframe of either 90 days or one year. Because they are probabilities, it isn’t really suitable for a single ward or specialty, but it’s very accurate for very large groups of patients. Next.

Clinical assessment reporting and tracking, CART-CL. It really doesn’t sound very specific by the title, but this is actually . . . was described by one study as an electronic health record tool to support VA cardiac catheterization laboratory quality improvement, a system for cath labs program. Another one called it a national clinical quality program for Veterans Affairs catheterization laboratories, so this is all about catheterization laboratories. It consists of 30 domain tables of cath lab data, and then it’s VINCI raw, so one thing I'd like to mention about VINCI raw, don’t shy away from VINCI raw. [Unintelligible 33:24].

Okay, next, patient aligned care team, PACT. This represents an index to assess patient-centered medical home model implementation. The index is based on a 2012 observation of more than 5.6 million Veterans, and the single purpose of the index is to measure patient satisfaction, rates of hospitalization and emergency department use, quality of care, staff burnout, and this is one of the few data providers that requires a data use agreement.

Brian Robison: Okay, thank you Jim. So now I’ll take over here and talk about first the DaVINCI domain. And so DaVINCI is actually a collaboration between the Department of Defense and the VINCI team to work on integrating Department of Defense data with data in the VA, and the Department of Defense is providing data from the military health system data repository, and this contains healthcare records and also demographics and service-related information from the Department of Defense, and the hope is to have the interagency collaboration and resource sharing and make this available to researchers.

The next domain is the equipment inventory. This is found in VINCI\_inventory. It contains information from the automated engineering management system, and it contains work orders and the inventory system for the different sites that automatically update every week, and so you can track different inventory and equipment as purchasing orders, all of that fun stuff as it gets updated.

And the next section we’ll get into is the SAS-based datasets. And so there are a number of datasets that are provided in a SAS format that are updated at different intervals, monthly, quarterly, annually, et cetera, and the VINCI team actually takes these and loads them into SQL server. We merge them, index them, and add in some fields that you’ll be able to join using SQL on. The first of these is the HERC, or the VA Health Economics Resource Center, and what this has is it is a way to determine the cost of care at the VA, to determine the cost effectiveness, evaluate efficiency, and it has the two datasets for the average cost datasets, which is basically an estimate of the cost of every VA health service, and so the estimate reflects expected costs given the nature of the service, and so they don’t really reflect differences by geographic location or other efficiency measures. And then the other one is the HERC V21 and Nosos, and these are risk scores that are calculated for all VA patients, and so these can be used to kind of estimate different risks for patients based on demographics and other diagnostic things.

The next is the MedSAS, or medical SAS datasets. These contain administrative data that involve patient care encounters for healthcare provided by the VHA, so this includes inpatient, outpatient care, and these files are extracted from the patient treatment file and then uploaded into SQL, and you can see some samples of the domain tables listed here.

Next is the planning systems support group, or PSSG, and this is the geocoded enroll files for a group of . . . basically it contains detailed geographic information, and this is broken down to one record per enrollee, and the VINCI team receives these SAS files and again we load them into the SQL server. We do not at this time, however, have access to the ADUSH, or Assistant Deputy Under Secretary for Health policy and planning files, so those are not included here.

Next is the VETSNET, or the Veterans Service Network Corporate Mini Master file. This is an extract from the Veteran Benefits Administration corporate dataset. So this includes information on disability ratings for both service and non-service-connected disabilities. It breaks down types, amounts, and dates of benefits paid. It provides info on the military service branch, entrance and release from active service dates, and non-compensation and pension benefits.

The vital status domain includes information such as date of death, the Medicare yearly enrollment indicators, date of last activity, and Veteran status flags, and this one is updated quarterly.

Then the VASQIP, or the VA Surgical Quality Improvement Program, is data provided through the National Surgery Office that looks at surgical outcome data including morbidity, mortality, and this provides breakdowns on the quality and processes and structures of surgical care at the VA.

So next I'll turn it over to Tori.

Tori Anglin: All right. Thank you, Brian. Okay, so today I'm going to be reviewing different services and supports available here at VINCI. VINCI data services staff provides support, education, outreach, data needs, project needs, prep to research, study setup, data access, cohort selection, provisioning, and along with other assistance.

So the first one I'm going to touch on is preliminary consulting and data needs assessment. Rarely does a study make it from conception to completion without questions about data availability, data attributes, and the data request process. We are here to answer your questions or find someone who can. By sending us an e-mail at VINCI@va.gov, we will create a ticket and assign it to one of our staff members that can help you with whatever needs you may have. We also have a link provided on the slide and at the very end for anybody who is new to VINCI. This page will walk you through different types of research. It will review the prep to research, IRB research, and different operation and quality improvement projects.

Prep to research allows your team to complete preliminary feasibility studies. As a PTR study, you would have access to VINCI workspace, CDW data such as the dim tables that were discussed earlier in this presentation, and your own database. This would allow your team to be able to use all the VINCI tools that we do provide. Once the study is approved by NDS for data access, VINCI data managers will help you locate your database, access the data, and provision your cohort, and this will all happen on our correspondence communication site.

All right, only approved project team members will have access to data. This is so VINCI can ensure the safety of the data. We also ensure the safety by storing it on either VINCI server or on local VA servers if your project is approved for that. In addition to this data storage, VINCI services include a cluster of servers set aside for tasks like data analysis, data processing, data extraction. This mean that VA researchers will have access to data and all the applications they need to select, transform, and analyze patient data in a central and secure location that’s acceptable from the VA Internet.

All right. There are three types of VINCI projects. When you’re going through your DART application, and along with your IRB, you’re going to need to select one of these locations. The first one would be VINCI workspace, which means that all of your work is done within a secure VINCI environment. Your study will have access to all the VINCI applications and tools provided. The second option would be to select to use your local VA server. This allows you to install your own software and assigns you a development workspace. The catch with this is that the data is not live but rather a snapshot or what we call read-only format. The third option is called VINCI download. It’s a combination of the two, so you can have a project in a live environment that you can work on, but then you can also download it to your local server when you want to do further analysis.

After a research project has data access approval, the VINCI data manager will help with creating your cohort. When creating your cohort, you have a couple different options. If you already have a list, patient identifiers like a list of socials, you can let the data manager know. We’ll provide you with a secure link that I’ll mention in the next slide and help you get your cohort set up within the VINCI workspace. The other option would be to just help you manually filter all the CDW data and get you down to the cohort you’re looking for based off your project’s inclusion and exclusion criteria.

So if your team decides to upload a cohort that does contain patient information, it’s protected, and so we don’t want to put it on a correspondence site, so the VINCI data manager will just need to know that you have the list. We’ll provide you with a secure upload tool. From there, we’ll put it into your VINCI workspace, and we’ll either help you connect it into your cohort or we’ll give it to you from there. Like I said, if you didn’t want to do that, we can help you with provisioning based off your inclusion and exclusion criteria. And then another option we do provide is if you ever wanted to . . . sorry, on the next slide . . . use external data, say you have a file, you can either upload it using the VINCI Central tool if it’s under two gigabytes, or most datasets are larger than that, so we would upload it for you. Another option was using external data. A lot of times data will have a data steward for different registries. You can contact the data steward, let them know you’re working in a VINCI space. They can get ahold of us, and we can help them move the data into VINCI so you can work with it in a safe environment.

All right, I'm just going to quickly touch on a couple different links that are on the VINCI Central page. So if you visit VINCI Central, there’s a tab called quick links on there. It will provide different links that we offer. The applications is another tab that we offer that provides other like the dim data viewer. Another tab that you can see on VINCI Central is user guide. User guide shows the guides for VINCI workspace, different software, OMOP, SAS. And the last link on this page is data sources. It’s also listed on VINCI Central and can link you to the CDW domain page and different VIReC documentation.

Lastly, some other helpful links that I would suggest all researchers look into is the VHA data portal. This is known by a lot of people as the one stop shop for data users, and that’s because it contains multiple sources in one single location. On this VHA data portal, it also has the links for VIReC, CDW, and BISL, and you can subscribe to the Listserv, which is a VA e-mail that is sent out to different VA researchers when they have questions and they want to share info. Those are all some very helpful sources. Once you get access into your VINCI workspace, this is just a sample script that you could use to search different domains for key words. In this example, we were searching the CDW work database for the schema name inpatient.

I'm sure plenty of you guys have questions. We’re going to take your questions. If you have further information needs, you can e-mail us at VINCI@va.gov. Now we will open it up to any questions you guys have.

Rob: Wonderful, thank you. We do have a couple of questions queued up. Audience members, if you would like to ask the team a question, you can go ahead and use the question pane in the GoToWebinar dashboard. Again, it’s one of those white areas with the gray lines dividing them, and you just click on the white triangle, and it will open it up. So I’ll just launch into the first question, which is are data tables the same as fact tables? And that one came in pretty early on.

Unknown speaker: Yes. Data tables and fact tables can be used interchangeably, but they are the same thing.

Rob: Great, thank you. This next one is the only other question we have right now, but I think given the nature of your presentation, there should be more coming in. Nevertheless, this person asks, if I want to look at number of encounter and number of procedures by a provider for both inpatient and outpatient, how would I look for that?

Unknown speaker: The question again was procedures and what?

Rob: Sure, I can repeat it.

Unknown speaker: Yep, thank you.

Rob: Yep, sure. If I want to look at number of encounter and number of procedures by a provider for both inpatient and outpatient, how would I look for that?

Unknown speaker: [Inaudible 48:02]

Unknown speaker: Providers [unintelligible 48:05]. Yeah, we’re all just kind of conferring across the table here. There are a lot of different ways you could do that. You could go with provider SID. As far as which table or tables you would use, first thing that comes to my mind would be OMOP because it conglomerates a lot of this information into encounters or concepts. It takes a lot of the different data and puts them into one spot or one group of information, so probably OMOP would be a good place to start, but there are other ways to do it as well. You could join some other tables using provider SID as well as things like in the outpatient data you could use a visit SID, which takes all the information for that patient’s particular visit. I’d start with OMOP probably.

Rob: Okay, thank you. This question just came in. Have the presenters had any interaction with the OHDSI collaborative?

Unknown speaker: No, but that is one of our goals is to implement some of the auto-feed tools and make that available, so that’s why we’re trying to integrate OMOP into VINCI.

Rob: Thank you. My study received DART approval on March 29, 2018. Why hasn’t a VINCI project folder been established for my study yet?

Unknown speaker: Okay. They can shoot us an e-mail at VINCI@va.gov, and we can figure out why or what’s happened in the process.

Rob: Thank you. This next question: Thank you for the first answer. Also, how can I link a particular addendum by an attending to the parent document? For example, how would I look for addendums to the SICU admission note by a specific provider?

Unknown speaker: There’s an addendum number that links back to the original note, so if you have the addendum . . . really I think you would probably start at the note. When you get to the note and it has an addendum, then you can look it up by the addendum number. You can also do the reverse. You can take an addendum and look it up by the parent number.

Rob: Wonderful, thank you. At this time we have no more pending questions. Audience members, if you have anything that you’ve been holding back, you don’t think it’s important enough, now’s your chance. Other than that, Kevin or team members, if you have any closing comments, I'd like to give you a chance to throw those out there now.

Kevin Malohi: No real comments, but thank you everyone for joining us, and, again, if you guys have any questions, you can always contact us at VINCI@va.gov. Thank you.

Rob: We had one more question come in while you were . . .

Unknown speaker: [Unintelligible 51:26]

Rob: We had one more question that came in, Kevin, is that okay?

Kevin Malohi: Yes, we still have time.

Rob: Okay. Could published OHDSI protocol serve as the basis for a new project with VINCI?

Kevin Malohi: So I’m not too familiar with the OMOP domain. However, we are giving a OMOP presentation next month, so you can definitely attend that presentation as well as send us that question again to VINCI@va.gov, and then we can follow up.

Rob: Thank you. Well, that was a fantastic segue to the next thing I'd like to say, which is that VINCI makes a presentation every month, and next month is on OMOP. So, Kevin, this person says thanks, exclamation, to your answer. And audience members, when I go ahead and close the session momentarily, you will receive a survey. Please go ahead and take a few moments to fill out that survey. We do count on you and your answers to our questions to continue to bring you high-quality Cyberseminars. Once again, Kevin and team, thank you for this very informative and important Cyberseminar. Have a good day everybody.

[ END OF AUDIO ]