Cyberseminar Transcript Date: 11/8/18 Series: VA Informatics and Computing Infrastructure Session: Identifying Acute Inpatient Stays within the CDW Inpatient Domain & Inpatient OMOP Visit Occurrence Transformation Presenter: Michael Matheny, MD, MS, MPH

*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)

Moderator: I’d like to introduce our speaker today. Dr. Michael Matheny is the associate director of VINCI at the National VA and associate professor of biomedical informatics at Vanderbilt University Medical Center. Michael, I’m sending you the pop-up now. Can I turn things over to you?

Dr. Michael Matheny: Yes, and thank you. Let’s move this over. Can you see my screen?

Moderator: Not yet, you have to click show my screen. There you go, yes, we can.

Dr. Michael Matheny: Okay, great. Well, thank you very much and I very much appreciate the opportunity in a Cyberseminar setting to talk a little bit about some of the data transformations and some of the considerations for acute inpatient stays in the VA. So, I’ll go ahead and get started. I’d like to get a little bit of a benchmark for the audience to get a sense of how familiar you are with these data. And so I’d like to ask a few questions if you wouldn’t mind.

Moderator: The first question is up Dr. Matheny. And we’ll give people some time so that they can provide their answers. First question being, how much have you used the VA OMOP data resource? And we have about 70% voted. Usually tops off around 80 so we’ll give people a few more moments.

[pause 1:28-1:35]

And things have slowed down so I’m going to go ahead and close the poll and share out the results to the audience. But I’ll tell you Dr. Matheny that in answer to the question how much have you used the VA OMOP data resource, 71% of your audience today say none, 5% say a little, 16% say some, 5% say a majority and 3% say all the time. Do you have anything you want to say about that or would you like me to just launch into the second?

Dr. Michael Matheny: No, I think that’s very helpful to get a sense of how familiar the audience is with OMOP. So I think I, it reassures me that I set up my background appropriately. Yeah, you can go ahead and go to the next slide.

Moderator: Okay, so that poll is up. Question being, do you or your team have a code transformation to use inpatient data in the VA written for current or prior projects? Yes or no. And answers are streaming in.

[pause 2:32-2:38]

About 50% of the audience has answered so far so we’ll give people a few more moments to give their answer.

[pause 2:43-2:53]

Things have leveled off so I’m going to go ahead and close the poll and share out the results. And read to you that 54% say yes and 46% say no. So now we’re back to your slides.

Dr. Michael Matheny: Great, thank you so much. So that also gives me very helpful information about the audience. So I’m hoping for those of you that do use inpatient data but don’t have any transformation or cleaningcode that one of the take-homes from this presentation hopefully will be a need to do some of that programming in order to get at what you’d like to get at. But I will present my case. So…

Moderator: Michael, I’m sorry to interrupt.

Dr. Michael Matheny: Yes.

Moderator: We’ve had a couple people ask if you could speak up. So maybe you could move your microphone a little bit closer to your mouth.

Dr. Michael Matheny: Sure, yeah. And I appreciate that feedback. I have a headset. Is this a better volume for the audience?

Moderator: It’s a little bit better. I don’t know if you can do anymore. Let’s just go with that.

Dr. Michael Matheny: So, I will try to speak as loudly as I can. Is that better?

Moderator: Yes sir.

Dr. Michael Matheny: Okay, great. So first I’d like to give a little bit of background on OMOP as a common data model and why it’s important to have an understanding of some of the common data models that are out there. Next, I’d like to give a little bit of an overview and an anchoring on some of the work that the CDW team has been doing on extracting data from VistA into the Corporate Data Warehouse for reuse. And then some of their recommendations for how to use the CDW data. And then I’d like to transition into how we chose to take a subset of that data and transform it with a series of logic steps and conventions into the OMOP transformation in order to make it reusable, or more easily reusable, for a specific set of use cases.

First, to give you a little bit of background on OMOP. So, there’s been a really a rise in the last five or ten years for the use and need for common data models and some of you may be familiar with these. There’s the Sentinel and the Mini-Sentinel initiatives that the FDA funded. There’s the i2b2 from the CTSA that were funded for a data model extracts from claims data and electronic health records data. There’s also PCORI-funded data research network called PCORnet. And so you can see over time there’s been a lot of money and efforts spent in trying to transform electronic health records as well as claims data into more easily digestible and computable forms. And it’s because of the old adage that really sometimes it takes 80 or 90% of a project’s effort to go from source data to transform the data into exactly the representation that you need for your use case. And so it also requires a great deal of knowledge about your source data systems, so it’s really hard. The farther you get away from actually using the clinical system and understanding how the data is collected to be able to use it for secondary use projects.

And so rationale for a common data model is for a set of experts to come in and to gather conventions from the experts in that area and then try to do a transform that applies, that is successfully addresses the most used case needs for the least amount of effort. And so there are a number of these tools that we mentioned. VINCI did an environmental scan and looked at a number of meta-analyses looking for how widely covered the data model is for the common source data that research projects like. Things like medications, laboratory, bar code med admin, administrative codes, things like that, as well as the least data fidelity loss. So anytime you transform data from one representation to another, you can lose some of the granularity or you can lose some of the data columns, data elements that travel along with that data. So you have to be aware of potential concerns when you’re doing that. The other reason why OMOP was chosen as the preferred common data model was there’s an open source development community that really has a lot of data visualization, analytic preparatory to research tools, as well as phenotyping, data quality and there’s more coming out all the time. And so our target was really to make sure the data transformation, the content mapping, quality assurance activities could be actually leveraged into a different CDM, not just OMOP. So if OMOP were to fall out of favor, we wanted to make sure that the effort expended would be easily transferrable to another data model if the community started preferring that such as PCORnet or i2b2 or other data models.

So this is an overview of some of the activities going on for OMOP transformation in the VA and how we think about it from a design perspective. Because some of the data in the VA has regulatory restrictions or requirements, there are actually multiple common data models that are aligned with each other being built. So there’s the one from the CDW that VINCI is working on, and some of the natural language processing tools that various informatics research groups are building, are being transformed into OMOP and pushed into that too. There’s another large project that was co-funded by the DoD and the VA to take in the Department of Defense data and transform that into the OMOP common data model. And then lastly, the VIReC resource center is leading a transformation, taking the CMS and Medicaid and Medicare data and transferring it to the OMOP common data model. You know I’ve been in a situation where I’ve been involved to some extent with each of these efforts and my native environment is the Corporate Data Warehouse. I’m a part-time primary care physician in the VA and practice in that environment and so I had a fairly decent degree of familiarity with the source data. But it’s been amazing to me working with the DoD data experts and contractors, just how little I knew and how much error I would have made in my transforms trying to do that from the source data. And then again, I’ve learned a tremendous amount from VIReC in trying to assist them in doing this Medicare and Medicaid transforms. It’s amazing the level of detail and source data knowledge you have to have in order to be able to transform these data in a reasonable manner. So I think having source experts doing some conventions and some data transformations is really quite helpful to the community.

So this is an overview of the data schema. I won’t go into too much detail. I just want to point everyone to the OHDSI website which is a repository for all of the data definitions, online tools, open-source tools. And then there’s also a VA Pulse site where all the VA-specific implementation logic code is available.

And so just to briefly touch here. So the way that OMOP transforms its data at a metalevel, and I don’t have time to go through the entire pieces, but there’s a control terminology or a standard data representation that each of the different domains has as its standards. So for drugs, it’s RxNorm. For procedures and conditions it’s SNOMED and for measurements such as laboratory tests usually it’s LOINC for labs and SNOMED for some other domains. And so it’s helpful to know what the target domain is for each of these in terms of being able to know how to use them. So to give you just one example would be for drugs, VA represents its drugs in VA product codes or VUIDs or NDC codes and both of those have a crosswalk to RxNorm which allows transformation. And then since you might have data, like Korea has its national data transformed into OMOP and then Australia has some of its data. So not all of the countries and not all the participating partners in OMOP actually have the same source vocabularies, but if everyone transforms into a standardized common representation you can write your code, execute on your local system and execute it somewhere else and have at least a moderate degree of confidence that you’re going to get some comparable answers. And there’s been some really high-profile publications that have come out by George Hripcsak and others recently that have highlighted the success of this type of approach.

So let’s dive down from the high-level view down and get into the nitty gritty. So the topic of the rest of the hour is going to really be focused around acute inpatient stays and how to look at that data and how to get some of that out.

So it’s a very common frame of analysis for observational cohort data really because of its increased granularity. You have a patient that is continuously within the hospital generating data. You have a lot of touch points. You have a lot of laboratory, a lot of vital data. You can be certain that medical care being delivered within that time frame is all captured in the sense of whatever is documentable is captured within the EHR. And so it has a high volume of research and operational use. There’s multiple operational data cubes that use this data, IPEC and other resources and so it’s a very important frame of reference.

So I just wanted to get a sense on the research side of how many VA publications [unintelligible 12:55] we’re handing out per year in this area, so I did a very rudimentary PubMed search just trying to get a sense of it by year. And as you see in this graph, you’re looking at between ten and 30 publications a year, which only scratches the surface of use. I’m considering all of the SAIL metrics and operational data cube and larger analyses that are being done for business intelligence and for quality improvement as well.

There are some challenges in using this data. A lot of this stems from the underlying source representation of VistA and CPRS and how that data are inputted into the system. Depending on which VistA files you access and how you look at it, you can end up with getting different answers. And so when I make the statement here that there’s potentially 11% by volume erroneous data in the inpatient domain, it’s really a statement that depending on which lens you use to look at the data it may have more or less error. And when I make that statement I do it in the context of a research retrospective project wanting to look at observational data over a period of years and wanting to specifically focus on acute inpatient stays rather than long-term nursing home stays or other sort of domiciliary stays. So take that comment in that context.

So there’s a lot of problems, and although these problems have been mitigated over the years, meaning there’s a lower rate of error in recent years than there were in distant years, there’s a lot of problems. As data are entered in the system sometimes someone was in a nursing home or in a domiciliary care in the VA and they’re brought into the hospital for care, they don’t stop the other record. So they leave the nursing home record active to keep the bed for them, they get brought into the acute stay which generates another record and then they end up with multiple records of overlapping time.

There are other problems where long term and acute care are recorded as sequential movements within a one larger record. So you might have a length of stay where if you didn’t really look under the hood you would think it was say 313 days when in reality it could even be three different acute inpatient stays of three days here, four days there, three days here along with rehab stays and nursing home stays and domiciliary stays. So, just trying to use that data as is without some untangling of it, you could end up with a certain error rate in your retrospective analytics.

And then sometimes you can actually have multiple records created for a single inpatient stay with potentially overlapping dates. And so then you actually have an additional complication on the outpatient side where there’s a lot of inpatient care that actually gets loaded into the outpatient domain. And this is a historical artifact from the fact that all the specialty consults really need to have their workload tracked and justified in order to be able to get their [unintelligible 16:03] and maintain their clinical services. But the system doesn’t have a good way of tracking that workload. So you end up with inpatient consults that get captured as outpatient visits and get workload credits through the outpatient visit sequence. So that creates some complexities in managing all of these data.

So I want to go ahead and just do a brief overview of the CDW Inpatient 3.0 domain. I first want to make some big key points. So one is CDW’s primary mandate is to extract the underlying source production data from the VistA and CPR systems and make it available to operational and research users across a tremendous set of use cases. The next thing I would say is that it has been really cool to read and look at all the work that the CDW team has done on the inpatient domain over the years. So, Richard Pham, Trinity Hall, Steven Anderson have gone through multiple domain iterations and really, really this is one of the most complex data domains in the system and it’s critically important because it’s an encounter anchor. So I would just direct you to all of their documentation and PowerPoint presentations and work because none of this transformation to OMOP would be possible without all of the complex work that they’ve been doing. So I want to make sure and make that point. The other key point here really is that, if you think about it, data warehousing is conceptually a little bit of a different task than trying to represent a common data model for specific category of use cases which is generally the cases in implementation such as OMOP, i2b2 or PCORnet. There are different tradeoffs and requirements for source data, target data representations, usability and considerations for handling and merging massive data volumes. I’ll just use one example before I move on which was, the CDW maintains everybody in a patient SID which is essentially a unique ID that concatenates the patient identifier at an individual site with the station. Then there’s an enterprise level of maintenance, so the patient ICN that is meant to disambiguate that. And I’m sure many of you have wondered well, why doesn’t the CDW create a single source of truth, one record for a patient so that you can get a date of birth, date of death, race, ethnicity without having to go through some transform to do that? But, if you think about it at an architectural level and a data management level at a database, managing 130 sites and then doing a cascading merger of records across billions of rows when sometimes those pointers change, is an extremely complex data task and one that is really not what a data warehouse is supposed to do. So I just, from an IT and informatics perspective, you have to be aware of what the systems are targeting. I felt that was important to go through that a little bit.

So let’s go into the content. I don’t know how many of you are familiar with ADT systems, but essentially this is called admit and discharge and transfer records. So when someone hits the hospital there’s an administrative layer where there’s a clerk tracking that data flow at every point in which there’s a change in the system. So when someone comes in from the outside, there’s someone there that tracks them, that documents where they came from and where they’re going to and who’s taking care of them, what the physical location and what the medical specialty or clinical service is that is taking care of them. And all of that gets date time stamped. Complexity comes in because there are multiple places in which this can be documented and multiple files in VistA that can store some of this data. And so it can get a little bit difficult in terms of trying to merge and get to single source of truth.

This is a slide from a presentation from CDW from Pham and Hall which just shows that for different target domains in CDW, you actually have to pull data from multiple different [unintelligible 20:13] and file sources. And again it’s beyond scope of this presentation to go through that. You can refer to their documentation for that. But I just, it’s quite difficult to get to a single source of truth.

So, the way that CDW built the data for the Corporate Data Warehouse was they pulled data from multiple VistA files, and I’ve listed them here but I won’t go into detail. But, in order to support all operational and research use cases, they provided all of the versions of these tracking systems. So provisional movement, inpatient transaction, patient transfer, specialty transfer, and so they had certain recommendations for what you should use when you’re wanting to know what’s the admit data? What the discharge date? What’s the specialty? What are the ward locations and things like that? And again, I point you to their documentation for that reference. OMOP really wants you to have a specific frame. So it wants you to have a single record for an inpatient stay. Because this is retrospective and not prospective or real time, that simplifies some of the design considerations, but there can’t be any overlapping records, so a patient is not supposed to be in two places at one time. We made the design choice to focus on inpatient and inpatient transfer which is one of the versions of the source of truth for this type of data. Not the only one.

I wanted to give you a little bit of an overview of this data. So you’ve got the two primary frames of data that we use to transform to OMOP, our inpatient and inpatient transfer. And so you can see here that the areas highlighted in green are how those two tables are linked together. They’re linked together through a patient identifier and through an inpatient record stay. And so another way to think about this is the, in the CDW the inpatient data table is a rollup of these individual ADT transfer sequences in an attempt to describe the situation. There’s a couple different key domains of data. So one is, where is the patient coming from? Which you can see is in the highlighted boxes in both the overall inpatient record and the patient transfer record. And then you have a where are they now? So, the difference between the patient transfer record and the inpatient record is the inpatient record has information about where they came from, what they were doing while they were in the hospital, and then where they went to after the end of that inpatient record. But the patient transfer record is just a state change. It’s a snapshot of a change in data. And so that’s why there is no what’s happening right now in the patient transfer record. And then where they were going to is happening in these fields. Where they’re discharged to, where they’re transferred to. What type of facility movement? What type of specialty in-service? What’s the physical location? All these things actually give you information that’s important to be able to disambiguate and try to track what is an acute inpatient stay.

And so, OMOP transformation. When we targeted the data, so there’s the encounter data for the inpatient domain and then there’s all the data that hangs off of that inpatient encounter. Things ranging all the way from barcode med admin to condition codes to procedure codes. So when you try to calculate an inpatient record with the era of the time associated with that inpatient record, you have to also take care or at least be cognizant of a huge amount of downstream data that is linked to that data. And so again, I just want to point your attention to the fact that the CMS place of service is the vocabulary that is used for OMOP and visits. It was originally a claims data vocabulary but then they added some electronic health record derived data elements to it that are relevant.

And this is what the OMOP table looks like. I’m trying to really walk a fine line in this presentation by not going too far into the weeds and I’m happy to touch base with groups afterwards by email or phone and talk about things in more detail. So you’ll see this presentation try to take a fine line approach to that. There are some standard fields above and then there are some custom fields below. Standard fields are the ones that every instance of OMOP is required to have and then the bottom ones were custom ones that the VA OMOP added in order have providence of the data to be able to track it back to the source and to be able to execute change management of data over time.

This is a breakdown of the fields in the table. So essentially you’ve got a pointer to the person, you’ve got a pointer to the provider of the inpatient stay or the outpatient stay, you’ve got a pointer to the location of care and then you’ve got a date time stamp from when the stay started and ended. And you also have a pointer to a concept that tells you whether or not this encounter is an acute inpatient stay or a nursing home stay or an ED stay, those sorts of things.

So this is an excerpt of the types of data. So if you’re using OMOP you’d essentially just focus in on acute inpatient stay or inpatient hospital and all these other categories you would be filtering out of your queries. And so there are some conceptual challenges in trying to transform the data from the source system and the CDW into the OMOP common data model. And really the biggest one is accurately separating acute inpatient stays from long-term care because of some of the issues that I mentioned to you earlier about how things were strung together serially or duplicated. And then we have to worry about how to handle attached downstream data such as administrative codes and laboratory tests.

So I’m going to take you through psuedologic for how we thought about doing this transformation. In order to prepare to do this we talked to multiple COINs and multiple research groups and we looked at some of the observational groups conventions and recommendations to try to get at this. And I think that this is focusing on a certain category of use cases, so I would love to get feedback from different groups about how close or far this sort of logic transformation is to the way that you like to use and manage the inpatient data. And so you’ll hear me talk a little bit more about that. But be thinking about how this is the same or different than what you’re used to because we’d love to have an ongoing dialog about this. So first we take all of the data in the inpatient and the patient transfer domain. We remove the exact duplicates of records. Then we remove records where the patient SID is not matching between the primary apparent inpatient table and the patient transfer table which happens to a very small amount of patients. Then we generate two rows of data from each inpatient record. So if you think back to what I was mentioning before, within the apparent inpatient record you have a sense of where the patient came from before they came in the hospital, where they are while they were in the hospital and then where they went to after they left the hospital. So those actually have to be broken down into two state changes. They have to be broken down into the state change of from outside the hospital to inside the hospital. And then a state change from at the end of the hospital stay you change state from being in the hospital to outside the hospital. So you can essentially normalize an inpatient record and split it into two rows of data. And then you add that into all of the transfer records which are additional state changes. And then you start merging that data.

There’s a lot of determinations that needed to be done and a lot of mappings and harmonizations. So first, there are a number of ways to determine where someone was based on a certain set of characteristics. So one, if you want to know what type of care they were being given, which is referred to as movement type. So you can either have intermediate care. You can have ASIH which is acute inpatient hospital. Ward, which is acute, usually hospital but not always. You have determination of what service they’re being transferred to and from whether that’s a medicine service or a nursing home service. And then you have a large set of Boolean rules that you have to develop in order to collapse essentially synonymous types of services and movement types and determine whether or not that movement type is likely to be acute or long-term stay. All of those rules are available in VA Pulse and from the code to the community.

So then we sequence all of data for each of the patient SID based on the date time stamp of that state change. We merge all of the sequences that have the same place of service type and then merge all of the subtypes that are deemed to be synonymous. Merge all long-term care such as nursing home types together and then merge domiciliary types together so you end up with increasing segments of time as these multiple stays that are actually the same type of stay from an OMOP perspective, get merged together.

Then the last stage is the most computationally complex. You analyze all the overlapping records that are date time stamp overlapping and you attempt to disambiguate those time segments. So you essentially look for acute and non-acute segments, then you can break them apart into inner records and outer records if you’re able to determine the acute segments. For acute overlapping segments, if they’re both acute and they’re overlapping in time, you just merge them. And if they’re both non-acute and time and overlapping you merge them. So again, there’s a level of granularity issue here so it might be that someone is domiciliary and then transferred to a nursing home. Since those are both non-acute for the purposes of this transformation they would be merged as a non-acute setting.

So, for source data there’s 13.6 million records as of 10/1/99 in CDW. And after all of those transformation steps there are 15,021,000 rows in the data. This does include both acute stay and long-term nursing home care. To give you a sense of the magnitude of change, of all those records the ones in which the admit and discharge date time stamps were unchanged from the original impatient record were only about 25%, but that overstates the case. If you look at the date stamps, in other words if you ignore the fact that maybe the record changed by four seconds when the actual patient movement record said it was four seconds different than the parent record and you just look at the dates, then the amount of unchanged records is much, much higher. It’s 12 million out of 15 million.

So some of the key issues here are that there are a lot of source records that had to be merged. So there are about 800,000 records that had to be merged. Some of them were just two records merged to one. Some were three to one, some were four to one. And there were also a fair degree of records that had to be split. And those were examples like, you had a nursing home visit and then they were admitted to the hospital for acute and then they were sent back out to the nursing home and all of that was one record. There were about 811,000 of those records that were split.

All the administrative and other data that are attached to that encounter, such as administrative codes, we made the design decision to retain the original source date stamp of the CDW record in order to maintain source data fidelity. And that way it also allows you to determine which one you want. Do you want the date time stamp of the ICD-10 code, or do you want to link that ICD-10 code to the inpatient stay that it’s associated with and use the discharge date of that stay? You can do either with the data representation in that format. And then, when the resulting encounter is not a one-to-one mapping with the associated data, meaning that you have a ICD-10 code or you have a procedure code and it could actually map to multiple different encounters because of the constitution, then the associated administrative data is left without a visit pointer because it’s not deterministically, you can’t determine it 100%.

There’s a couple summary points. So inpatient data from the source really are not useable for most cases without significant cleaning and CDW gives recommendations for how they recommend that you clean it. And for certain use cases IPEC has a way that it does its cleaning and then OMOP has its cleaning version. So I would recommend that you look at these multiple versions to decide what works best for your use case. So we’ve done an initial quality assurance effort, but we welcome additional validation and collaboration and on the next slide I’ll get into that in more detail and I would just say that this is a very complex logic. It’s over 4,000 lines of SQL code. You could represent this in Python or R or C# or whatever you wanted to, but it takes a lot of execution logic in order to synthesize this down. Having to redo this for every research project, and then have to represent the intermediate data products for every research project is pretty a substantial effort.

So in the big picture thing, summary. If you’re using OMOP and you want to use the inpatient data in OMOP, usually one to three database joins for most applications can get you to a fairly clean representation of the inpatient data where you can start adding on your covariates for use in an inpatient frame. Again, as I mentioned before, it’s a lot of processing to get all this data represented for acute inpatient stay use case. And so having it processed and then the processed data useable for research projects allows you to free up space in your research database for other intermediate table products that you need for your use case. And it’s reproducible. IPEC has its standard version and so you use IPEC and you get that standard version and so OMOP is the same way. It’s reproduceable within the context of OMOP if you choose to use that transformation logic. And it’s transparent. VA Pulse has all the documentation and details about the logic and the SQL code itself are available to users so you can see the conventions and the concatenations and potential things that are with or against your potential use case. So I do want to emphasize the limitation here which is that it’s an enforced convention so based on the [unintelligible 36:10] documentation and expert consultation we transform the data using one of the possible paths that the source data allows, but there are different ground truths for different use cases, so it may not be applicable to your use case. We hope that it’s…and so the last big point is consensus building. So we’re trying to make this useable for the majority of research use cases. So we want feedback. We want criticism of the transform. We want to know when it’s not working well for your use case or if there are errors or gaps so that we can make an attempt to try to maximize the transform for the widest possible set of use cases for the research users.

And lastly, I just really want to acknowledge a number of folks. So I want to acknowledge Scott. VINCI has been supporting this work for a number of years. There’s also been PCORI funding and DoD funding and NIH funding to help some of this move forward. There’s been a lot of people, Kristine Lynch has been leading the quality assurance efforts. A number of developers have been working on the ETL and documentation and quality assurance. And we have a help desk that is available for anyone to ask questions for why something is done or how to do it that Liz Hanchrow leads. It’s been a very large team effort and very much appreciate all the work that everyone has done on this.

And so I purposefully left some time for questions. So, I’m happy to transition over to that point.

Moderator: Thank you Dr. Matheny. At this time we don’t have any pending questions which is strange for a VINCI Cyberseminar. So I anticipate some will be coming in. Audience members, in order to ask us a question, ask Dr. Matheny a question, please use the questions pane in the go-to webinar dashboard. As I explained at the beginning of the session, that white control panel that came up when you first joined the session. We did get one person just now saying thank you for making this phenomenal resource available to VA researchers. Not a question.

Dr. Michael Matheny: Very welcome. We hope that it gets use and that we can continue to improve it for everyone. I’m not seeing any questions in my chat panes. Is it not visible to me?

Moderator: It’s not. Apparently go-to webinar decides that it’s too distracting for the presenters. One person is writing in asking, can you use OMOP to compute bed days of care? And then in parentheses they write BDOC. I guess that’s bed days of care.

Dr. Michael Matheny: So, for the acute inpatient stays, it is a computed admit, and discharge date time stamp down to the second using the patient transfer record version of the truth. So you can compute bed days of care. So that if you’re looking for per-ward, like if you need to know how many days they’re in the ICU vs. how many days they’re in floor three vs. how many days they’re in the step down, this version of OMOP does not have that. Although they’ve just issued an update to the data model that has both the visit occurrence and a visit detail which is actually intended on tracking every change state in the process. And so future versions of the transform will include, you know hey, they spent a day and a half on floor three north and then two days in the med ICU and then a day and half in this, which will allow a pretty granular computation of time for various locations.

Moderator: Thank you. Another person is asking, what is the name of the Pulse page?

Dr. Michael Matheny: Actually I, let me see if I can find that real quick. So Liz, if you’re on, post that as a question. Liz is our help desk person and see if I can get to it real quick. [pause 40:24-40:29]. Should have had that in my slide. I apologize to the group. [pause 40:31-40:39]. Can everyone still see my slide? I mean, my desktop?

Moderator: Yes, sir.

Dr. Michael Matheny: All right. So that is the link.

Moderator: Okay, let’s see. It’s a little small on my screen. I can’t quite read it.

Dr. Michael Matheny: Can I post it in the chat window?

Moderator: You can, yes. And then I can put it in the answer to the question. Oh, we do have another question.

Dr. Michael Matheny: Go ahead, sorry.

Moderator: Okay. You mentioned the SQL code used to create OMOP tables is available. Where?

Dr. Michael Matheny: It’s available on request. So just ask [vinci@va.gov](mailto:vinci@va.gov) and we’ll send you the code. The reason why we don’t have it on the Pulse is because there’s a large amount of code at the beginning and the end of each script that has nothing to do with the transformation logic. It has to do with handling all of the change elements. So when you have given data that changes in VistA that then changes in CDW that then we have to account for, sometimes that results in one correction. Sometimes that cascades into ten and 20 corrections depending on what the data element is changed. And so that requires a certain amount of code bloat, essentially, that doesn’t have anything to do with the logic. We’ve had it on our plate to cut out that and just provide just the transformation logic in Pulse and we’ve never been able to get the bandwidth necessary to do that. So right now we’re just offering the code with all of that stuff at the beginning and the end on request. So just send us an email and Liz can package it up and send it to you.

Moderator: Thank you. What would be the approach for determining whether people received mental health care, parentheses, and what kind, during their inpatient stay?

Dr. Michael Matheny: That’s a good question. Let’s see, you would go to the visit occurrence table and then you would filter on the acute inpatient hospital stay. I think it’s like 920…you go to the concept table and search for the domain that is visit type. And then that gives you all the visit types. You pick out the inpatient hospital one from that. Then you join on that. That gives you all the acute inpatient stays in OMOP. And then you would separately go to your administrative codes, depending on your definition code list, and you would find…Probably what I would do is I would find all the administrative codes that are inpatient type. You’d have to decide if you just wanted the primary diagnosis code, or whether you wanted primary or secondary and then you could pull all of those codes for those patients which would give you the date time stamps and the links back to the visit encounter or the inpatient stay. So you could get your cohort that way.

Moderator: Thank you. Let me see if there’s another question that came in, I can’t…[pause 44:00-44:05]. No, there are no pending questions at this time. Perhaps if you have something that you wanted to go into with a little bit more detail, or if you have some closing comments, we could do that until other people have questions.

Dr. Michael Matheny: I guess I would actually…. No, I don’t have any additional comments. I would just encourage everyone to ask questions and use the help service. And it’s fairly easy to get access to the OMOP resource if you don’t already have it. It’s just an extra check box in DART. Basically, it’s a one-liner addition to your IRB. We are going to be using OMOP as one of the data sources. It’s a one check box addition to DART. And then the views get provisioned into your research database. And basically, you sit those alongside of a CDW so you use them both together.

Moderator: Great, thank you. For people who wanted the answer regarding the VA Pulse page, I did send it to everybody through the chat. I’m sorry, through the questions. But I can read off. It’s your normal VA Pulse path which is <http://www.vapulse.net/groups/vinci/OMOP/users/group>. So it’s VINCI, OMOP, user, group. Should be able to find it if you do a search in VA Pulse. Dr. Matheny, thank you for preparing and presenting this session today. As always, we appreciate all the answers that VINCI people give. And with that, I’ll say thanks to the audience members and if you would please fill out the survey, the short survey that comes up when I close the session momentarily. Again, Michael thank you very much and have a good day.

Dr. Michael Matheny: Yeah, I want to thank everyone for your time today.

Moderator: Thanks.

[END OF AUDIO]