Cyber Seminar Transcript  
Date: 06/29/2015  
Series: VIREC Database and Methods  
Session: Veterans Pharmacy Data  
Presenter: Walid Gellad and Xinhua Zhao  
 *This is an unedited transcript of this session. As such, it may contain omissions or errors due to sound quality or misinterpretation. For clarification or verification of any points in the transcript, please refer to the audio version posted at www.hsrd.research.va.gov/cyberseminars/catalog-archive.cfm.*

Moderator: Today’s speakers are the Walid Gellad and Xinhua Zhao. Dr. Gellad is an associate professor at the University of Pittsburgh School of Medicine and Public Health. His research focuses on physician’s prescribing practices and on policy issues affecting access to medication for patients. He is currently studying the overlap in prescription use among veterans cared for in multiple health systems. Dr. Zhao is a research health scientist at the VA Pittsburgh Healthcare System Center for Health Equity Research and Promotions. She received her doctorate in Biostatistics from the University of Pittsburgh Graduate School of Public Health and has been working VA Healthcare research projects including medication prescribing and safety for over five years. Questions will be monitored during the talk and will be presented to the speakers at the end of the session. I am pleased to welcome today’s first speaker, Dr. Gellad.

Walid Gellad M.D.: All right, thank you. I hope everyone can hear us. Thanks to VIReC for this invitation. I also want to thank Kevin Stroupe, who has given a version of this talk for the past several years. A lot of these slides we took from him and amended some of them and added some others. I am appreciative of his help on this. We are going to aim for 35 to 40 minutes of lecturing and then hopefully we have plenty of time for questions. We will get started. We have a couple of poll questions to begin. The first poll question, it says, “I am interested in VA data primarily due to my role as” blank. The options are research investigator, data manager, project coordinator, program specialist or analyst, or other.

Heidi: If your response is other, please type what that other role is in the question screen. I will go through those as we are going through the poll results. I will give you all just a few more moments before I close the poll question out. The responses are coming in well. It looks like things are slowing down. Some of the responses that we are seeing are 31 percent saying research investigator, 17 percent data manger, four percent project coordinator, 31 percent program specialist analysis, 17 percent other. There we are seeing healthcare inspector and dissemination and communication. Thank you, everyone.

Walid Gellad M.D.: Great okay, we have two more. Just get your fingers ready. The other one is “Have you ever used VA Pharmacy Data? Yes or No.”

Heidi: Responses are coming in nicely. We will give you all just a few more seconds before we close this one out here. Here we go. We are seeing 46 percent saying yes and 54 percent saying no. Thank you, everyone.

Walid Gellad M.D.: Great, okay, that is good to hear. This will be a lot of useful information for folks who have never used Pharmacy Data before. One more poll question is, “How would you rate your overall knowledge of VA Pharmacy Data?” On a scale from one being never used to five being used frequently or very familiar.

Heidi: Once again, we will give everyone just a few more moments to fill that out. Responses are coming in nicely. I looks like things are slowing down here. What we are seeing is 45 percent saying they have never VA Pharmacy Data, 22 percent rating at a two, 16 percent at three, 14 percent at four, and 3 percent they used frequently and are very familiar with it. Thank you, everyone.

Walid Gellad M.D.: Great, okay so the ones who answered five are not allowed to ask any questions and the ones who have never used can ask all the questions. We will leave questions to the end, but let me get started. We have a number of objectives. I am going to start and then I will pass it off to Xinhua and then it will come back to me. First, we are going to just give a broad introduction to how outpatient pharmacy data has been used in VA studies. We will add some details later in the presentation. Then Xinhua will come on and give an overview of VA and Medicare pharmacy databases. Then I will come back and talk about how to find information in the VA Medicare pharmacy databases and focusing specifically on how you identify drugs of interest and a few words also about costs. Then finally, I will talk through some examples of VA studies that have used the VA pharmacy databases and where to go for more help. Then we will finish it up with questions.

All right first, it is a brief introduction, a broad overview of how patient pharmacy data has been used in VA studies. I will go through three studies that we will talk in more detail later. The first study used the pharmacy data to look at trends in medication use. This is from Elizabeth Tarlov and Kevin Stroupe from a few years ago. It used pharmacy data to look at trends in anemia management in lung and colon cancer patients in the VA. Looking at trends over time and medication use is a very common way in which pharmacy data is used both in VA and Part D data. We will go into more detail on this later.

Pharmacy data has also been used in VA studies for cohort identification. This is just one example. This is from investigators from Houston. This is also from a few years ago. Here, pharmacy data is used for cohort identification. The goal was to identify rheumatoid arthritis patients and then use different algorithms, some of which included the use of pharmacy data. We will go into this in more detail later.

Then finally, this is a study of ours from earlier this year. In this case, pharmacy data is used both for cohort identification and we looked at utilization. We use VA PBM and Medicare Part D data for both of those issues for cohort identification and to assess the use of diabetes medications and test strips. This is really a true example of using pharmacy data on veterans from both the VA and the Medicare simultaneously. This will be a nice example.

There are many other really well done studies using pharmacy data, comparative effectiveness studies, studies looking at the quality of pharmacotherapy. These are just a few examples of many studies. We will focus on these three later on the seminar after we give an introduction to the data sets.

There I am going to turn it over to Xinhua and then I will come back. I may, Xinhua knows this, I may jump in a bit during of her portion, but otherwise she is going to talk and give an overview of the VA and Medicare pharmacy database.

Xinhua Xhao, Ph.D.: Okay, they are both local and national pharmacy data sources. The local one is VistA data. All of the pharmacy data entries are processed and stored in VistA. There are several national pharmacy data sources. The first source is PBM, pharmacy benefit and management data. The second one is DSS and NDE, Decision Support System, National Data Extract Pharmacy Datasets. DSS has now been renamed as MCA, Managerial Cost Accounting system. The third one is CDW pharmacy data, Corporate Data Warehouse pharmacy data. Medicare Part D data are from Part D Slim File, \_\_\_\_\_ [00:07:57] patient and pharmacy data. They have put important key pharmacy data sources, DDS product table, and National Drug File. These profiles are not personal level files. They are summary data.

Here is a poll question, “Which national sources of pharmacy data have you used in the past?” DSS NDE Pharmacy Data, PBM pharmacy data, CDW pharmacy data, and Part D Slim File, none. Please make your choice.

Heidi: We will give everyone just a few more moments to fill that out before we go through the results here. You do have the option if you have used more than one that you can select all that apply. It looks like things are slowing down here so I will close the poll question out. What we are seeing is 21 percent saying DSS NDE pharmacy data, 22 percent saying PBM pharmacy data, 33 percent saying CDW pharmacy data, two percent saying Part D Slim File, and 49 percent none. Thank you everyone for participating.

Xinhua Xhao, Ph.D.: Okay, great and now let us go through the full national data sources one by one. The first one PBM pharmacy data. PBM pharmacy outpatient data available from fiscal year 1999, inpatient data available from fiscal year 2003. The source of PBM pharmacy data is from local VistA data. The PBM pharmacy data includes records for both inpatient and outpatient prescriptions from VA pharmacies or CMOP, Consolidated Mail Outpatient Pharmacy. All CMOP prescriptions should be mailed to the patients. PBM pharmacy data is housed by PBM and available through custom extracts. PBM data includes the following key variables: VA product name, VA class, generic drug name, directions of use, dispensing unit, day of supply, total quantity dispensed, price per dispensing unit, total drug cost, release date, national drug code, NDC code, and CMOP indicator. From here, you can calculate the drug supply start data by the release data and day of supply. You can also calculate the daily dosage by the total quantity dispensed divided by day of supply. Dr. Gellad is going to talk about the example data so you can have an idea of what the actual data looks like.

Walid Gellad, M.D.: Yes, I think it is worth just spending a little bit of time here just to look at the specific example, especially for those who have not used PBM data. In this case, you can see the VA product name is diclofenac sodium. The VA product name is standardized across the VA and includes the generic name, which is diclofenac. It also includes salt, which is sodium, the dose 50 mg, and the tab which is the form. Then this is enteric coated. There is a lot of information included in the VA product. You can see the VA class. In this case it is MS102. Then a unique aspect of PBM data which we will talk about, you have the directions of use, two lines below the sig. It says take one 50 mg tab twice a day to relieve pain. Then you can see that the dispensing units, these are tabs. It is a 21-day supply so there is a total quantity of 42. Price per unit, the price per tab is 15 cents. You can calculate the total drug cost of six dollars, 34 cents. Then of course, it has the NDC and the fact that it is released by the CMOP. I think this is a nice example of what you can find in the PBM data.

Xinhua Xhao, Ph.D.: Now you have an idea of PBM pharmacy data. Let us look at the DSS NDE pharmacy data. Data available from fiscal years 2005, the same as PBM pharmacy data. It includes records from both inpatient and outpatient prescriptions from VA pharmacies or CMOP. The source of pharmacy data is the local VistA data as well. DSS pharmacy data is housed by CDW from fiscal year 2005 and available through custom extracts.

CDW Pharmacy Data is available from fiscal year 2000. The source of CDW pharmacy data is VistA data as well. CDW has two pharmacy production domains including both outpatient pharmacy data and inpatient pharmacy data. Inpatient pharmacy data is called Bar Code Medication Administration, BCMA. CDW outpatient pharmacy data is data from the PBM and DSS from the data because it is on one table. It is not just in one data field. Those tables are carved into two groups. The first group is outpatient pharmacy fact tables. Those are prescription level data and include four data fields here: prescription field, prescription fill field, sig field, and medication instruction field, data fields. You can extract the information from different data fields and link them together by a common identifier variable. It is called the RxOutPatSID variable. Please note in those prescription level data fields, some variables are not coded as the actual data. Instead, there are just some simple codes. You need to link the prescription level data to dimension tables that matches data tables to get the actual data information. For example, if you need dosage form information, you need to link the dimension table dosage form to the prescription level data with an identifier called DosageFormSID. Each dimension table has its own identifier to link to the prescription level data. Here we can see CDW pharmacy data includes lots of variable information, but also involves the actual data manipulation to link between different table fields.

CDW internet and website provides detailed information about those data fields. You can search for the domain lay out, data contents and sample data sets here to get an idea of what the actual data looks like and to look for the information you need for research.

Now let us compare the three VA national data sources we just talked about, PBM pharmacy data, DSS pharmacy data, and CDW pharmacy data. First, those data are available from different years. Second, DSS pharmacy data includes the cost information. PBM and CDW pharmacy data only include very limited cost information. If you are interested in cost data, maybe you want to get that for your research. The third difference is direction of use including scheduled dosage and dispensing unit are not available in DSS pharmacy data but are available in PBM and CDW pharmacy data. The last one is National Drug code, NDC code is available at the three data sites, but in DSS data, there is not a separate variable for NDC. It is contained in the field FEED\_KEY.

Regarding the pharmacy data comparable between those three national data fields, VIReC compared the DSS and PBM pharmacy data on outpatient prescriptions for a cohort and found that nearly all of the same prescription appeared in both files with a discrepancy rate of only 1.5 percent. This high consistency rate between the data sources are not surprising because all of the data national sources extract their original data form VistA. If you are interested in more information, you can go to the link here to look for more information.

All right, we just talked the VA pharmacy data sources. Now let us look at Medicare Part D data. In 2012, 40 percent of VA and Medicare enrolled veterans are enrolled in Medicare Part D. This means if your study cohort includes seniors over 65 years or older, 40 percent of those patients are likely getting some medications from Part D. For complete list of the drug information, you may want to consider using both VA and Medicare Part D data for your research. Part D data is different from Medicare Part A and B data. Claims for drugs are paid by insurance companies, not by the Center for Medicare and Medicaid Services. Insurance companies submit data to CMS on all prescriptions filled. There are several Part D data fields including PDE, prescription drug event files, characteristics files including information of data fields for plan, formulary, pharmacy, and prescriber. The Slim File is a subset of PDE data, like a mini version of PDE data.

If you are interested in Medicare Part D enrollment status, you need to find the information in Medicare Enrollment Files. There are two important variables there. The first one is number of months of Part D coverage. The second one is type of Part D plan. Part D Slim File extracts the key variables from the full PDE data from the Medicare Prescription Drug Event data. The Slim File data available from VIReC as custom extracts. Currently, in VIReC only have the 2006-2011 Part D data available. There is about a three to four years’ time lag here. You can see here Medicare Part D data is stored by current year, not like VA data, which is stored in fiscal year. If you merge the two data copies from Medicare and VA, you want to pay attention to the time frame. For example, if you are interested in fiscal year 2009 data, for Part D data, you need to use both calendar year 2008 and 2009 to get the fiscal 2009-year data.

Part D Slim File includes key variables from the PDE files. Most the variables are \_\_\_\_\_ [00:19:48] to pharmacy data. We can see here Part D not only includes the drug generic name, but also includes the drug brand name, but it does not include the sig variable such as the schedule and the unit dosage dispensed. We also give example data here. You can have an idea of what the actual data looks like.

For the patient level pharmacy data we just talked about, there are two important pharmacy data tables, DSS product table and the National Drug File. Those are not personal level fields, but they include very important information like VA drug dictionary. You can go there to search for the information of VA product names, VA drug class, and National Drug Code. Those two fields are available on the DSS intranet website and the PBM intranet website. Next, I will hand over to Dr. Gellad to talk about how to find the information in VA and Medicare pharmacy data.

Walid Gellad M.D.: All right, thanks Xinhua. We can talk more in the question and answer session about the Part D data and the relationship between the Slim File and other files and anything that comes to mind. Make sure you send any questions that are of interest.

Let me talk about finding information in the VA and pharmacy databases and focusing really on identifying drug and talking a little bit about costs. How do you search for medications of interest in these data sets? Let us say, for example, you want to look for all drugs for diabetes. Where would you start? Most simply option one, you could search by drug generic names or medication class. If you knew all the generic names for the diabetes drugs, you could simply search based on those generic names. The medication class is available in VA data, but it is not available in the Part D Slim File and Part D data. In VA, the medication class is HS501 and 502 are the diabetes drugs. You can search VA data using those classes. Again, there are nuances so those classes sometimes include drugs you do not want, for example glucose tabs or even insulin syringes.

You can also search by NDC code, or National Drug Code. For that, you need to use Medi-Span, First Databank or some other database to obtain the NDC codes. I would say it is not ideal to use NDC codes in VA if you do not have to because of the way that the VA standardizes the VA product name. We have used it before. We have used it, for example, to identify specific brand and generic versions of the same drug in VA data.

Let me just talk a little bit more about NDC codes. NDC is a unique ten or 11 digit, three segment number, which identifies the labeler, the product, and the commercial package size. In the parenthesis, here are the number of digits in each segment. The first segment is four to five digits. The second segment is four digits. The third segment is two digits. Here is an example of the ten-digit NDC code for a 100-count bottle Prozac 20 mg. You can see here the label or code in this case is actually a four, four, two format. The four digits here indicate the labeler. These four indicates the product. This product is Prozac capsules 20 mg. The package code here, the two-digit code indicates that it is the 100-count bottle. Together, these tell you the specific drug, product, and package size. VA PBM, the DSS, and Part D all use actually 11 digit NDC which would be a five, four, two arrangement. In that case, the leading zeros are added to the second that is too short. In this case, the leading zero would be added in the first segment. That is NDC code. NDC codes are available from databases like Medi-Span or First Databank or Red Book.

Now what about costs? Where can you find cost variables? In the DSS and the PBM, what is formally called DSS contains different cost variables. In PBM, it is quite simple. The cost that is included is the cost of the drug product from the supplier. It is simply the ingredient costs. From DSS, there are other variables, which are actually quite confusing. If you plan to use these variables, I highly recommend going to the HERC technical report, which is including in the VIReC user guide. That is really what I would definitely read if you have interest in using these variables. The three variables are dispensing cost. That is a direct pharmacist labor costs for dispensing the prescription. If it is a mail prescription, it also the mailing costs. If it is dispensed at the window, there is no mailing cost included.

Supply cost is the second variable. It is actually the cost of the supplies used in preparing the prescription like bottles and labels. It also actually includes the cost of the drug did not put here, but we should have. It is the cost of the drug plus the cost of the supplies. However, if it is a FEMA dispensed prescription, it is actually on the acquisition cost of the drug. These are the kinds of nuances that make it very confusing but HERC has a nice report on this.

Finally, the actual costs, the last one, is the drug product cost, again plus the cost of the supplies tike bottles and labels. That includes supply cost. Plus indirect costs and overhead. If you wanted the total cost for the prescription, it is the actual cost plus the dispense cost. It will give you the cost associated with that prescription. Those are in the DSS file.

What about in the Medicare file, specifically in the Part D Slim File? There are two cost variables included. The first is Patient Pay Amount. That is the amount the patient pays for the medication, which is straightforward. Then there is gross drug costs, which is a derived variable. It is the sum of all of these issues that we listed, the ingredient costs paid, the dispensing fee paid, total amount attributed to sales tax, and a vaccine administration fee, if that exists. Those are the low cost variables from part. D. Those are two examples of how you find drugs and how you look at cost. I am going to through a number of examples now, three examples that I talked about already, of VA studies that have used these pharmacy databases. These are just three examples of many actually. Here is a listing of some of the ways in which pharmacy data has been used in outpatient studies, it is used for cohort identification like we talked about. Can pharmacy data be used to identify specific groups of patients? You can look at medication utilization. Does a policy change impact medication use? These are very relevant for Part D. You can look at healthcare quality. Are patients being prescribed medications according to the quality measures? We can look at medication adherence and exposure to specific medications or medication class. For example, are specific drugs associated with better or worse outcome than other drugs? In some cases, there also have been studies combining outpatient and pharmacy data to identify events. Can you identify acute exacerbations of COPD with outpatient and prescription data? The answer being yes. That is nice example of combining those two data sources. One can also assess comorbidity or case mix with medication data.

We are going to talk just about these three examples. The first one, again, is from Tarlov and Kevin Stroupe looking at trends in medication use. The objective in this particular study was to examine erythropoiesis stimulating agents in lung and colon cancer patients getting chemotherapy. ESAs are drugs used to treat anemia in cancer. There was a black box warning that was put forward by the FDA for these drugs in 2007. There were increasing restrictions on their use during that time period. The point of the study was to look at trends related to those policies.

Here is the study format. There was a pre-period before the black box warning within March of 2007. Then there was a post period afterwards. Pharmacy data was used to examine whether ESA use was different before or after the black box warning and also looking at trends in ESA use over time. The source of pharmacy data here was the PBM database to identify ESA. In this specific case they used NDC, NDC codes to identify the ESA in the PBM database. A quick note actually and an important lesson here, not only did they use NDC codes, but they also actually used CPP and HCPCS codes, HCPCS codes from inpatient and outpatient counter data because these particular drugs can be given in the office setting. There are other drugs like that, most notably methadone, for example, that is administered in substance abuse clinics. That is just a quick lesson. The ESAs were not only examined by NDC code, but also because of provision from physician or provider offices for CPT codes and HCPCS codes. What did they find? These are the main results. These are adjusted odds ratios of ESA treatment and their variable pre-post although in this case, this is comparing the post period to the pre period. They found in the left column for lung cancer the odds of ESA used to create 65 percent in the post period. They also plotted, predicted probably from their model, which is what this is. They hit probabilities over time and the red is lung cancer and the black is colon cancer. You can see that the probability of ESA use started to decline actually for both groups before the black box warnings were issues around this time. It is nice example of looking at medication use over time. That is the first study.

The second study used medication use for cohort identification. The objective here was to assess the utility of diagnostic algorithms including prescriptions to identify rheumatoid arthritis patients in VA database. The study sample consisted of patients having two outpatients visits with RA codes, at least two months apart at the Houston VA between 1999 and 2009. The study tested algorithms to identify RA through these databases in a number of ways. First, most simply, was just the presence of at least two ICD-9 codes at least six months apart. They added the use of a disease modifying anti-rheumatic drug or DMARD for at least 180 days as one of the criterion for identifying rheumatoid arthritis patients. Then subsequently they required one of the ICD-9 codes for rheumatoid arthritis to be from a rheumatologist. Those are the different algorithms that we used to identify RA. They validated these diagnoses using chart review, which was the gold standard in this case. They evaluated the positive predictive value of the diagnostic algorithms using the chart review. In this case, the pharmacy data came from PBM. We are kind of PBM heavy here, but it is what it is. Here are the major results from the table. All of this information is not important. Really, what you can see in these columns, this is the positive predictive value. This is the sensitivity. This is the specificity of their different algorithms. The top half is an algorithm without including whether or not individuals use drugs. The bottom half includes individuals only using DMARD therapy. That was really the main finding here is that the positive predictive value was highest and the algorithm included the use of the DMARD to identify patients with rheumatoid arthritis. In fact, the highest positive predictive value included use of a DMARD and use RA codes specifically from a rheumatologist. You can see positive predictive value of 91.4 percent. Obviously, sensitivity suffered in that case even though specificity is very high and there is no RFD curve, but it is a nice example of looking at positive rates of value and using drugs to identify condition.

All right, now the last example, this is our study from earlier this year, which the goal was to examine their patterns of test strip receipt of among older veterans with diabetes and determine whether receiving strips from dual health systems from both VA and Medicare was associated with overuse. The design here was cross sectional. It was a cross sectional retrospective cohort. We used National VA administrative data linked to Medicare Parts A, B, and D claims for fiscal years 2008 to 2009. We had roughly 364,000 community dwelling veterans aged greater than 65 with diabetes. We used the VA healthcare system and received test strips fiscal year 2009. The source of the pharmacy data here is VA PBM and Part D slim files. They were Part D files. We actually used VA Medicare data for everything in this case, but for cohort identification and comorbidity adjustment, the medication use and test strip use, which I will show.

How was pharmacy data used? First, we used PBM data to search for diabetes medications by VA drug class and by generic names, which is what I talked about before. We combined that with diagnosis codes from MedSAS to define the study cohort of Type II diabetes patients. We combined medications use and diagnosis codes. We used PBM data and the Part D files. Here, again, Xinhua mentioned the need to sync calendar year with fiscal year data. We used calendar year 2008 and 2009, Part D files to classify patients based on the type of diabetes medications they used. In the Medicare Part D, we used NDC codes to identify these drugs that we obtained from MetaSpan. We Identified individuals in the cohort who are no meds, those who are on oral medications only, and those who are on long acting insulin, and those who are on short acting insulin. I will show you why later. It is in order to identify who might be overusing test strips. We identified their medication use.

Then we used PBM data to search for and quantify test strips, which are actually dispensed medications in the PBM file. In Medicare, we used the DME file, the durable medical equipment file to search for and quantify test strips in Medicare, which are not actually in Part D data, but in this case are in DME files.

What did we find overall among this cohort of test strip users, 71.6% received strips from the VA only, 22.8% from Medicare only and 5.6% or around 20,000 veterans received test strips from both the VA and Medicare. Let me show you some of the figures from the study. Here, now, is a figure of the median number of test strips. That is what is on the Y axis, the median number of test strips dispensed in a year. These are the different medication classes that we needed to identify both with Part D data and with VA data. Those individuals who are no medication, those who are on oral medications only, those who are on long acting insulin, and then any short acting insulin. If you focus here, what you see is among those on no diabetes medication, individuals who received their test strips from VA only received a median of 100 during the year compared to those who received their test strips from Medicare only or from both VA and Medicare who received around four times as many test strips in the year. You can see the relationship between VA only and then dual use across each of the medication classes. In doing this, we actually defined potentially excessive test strip use, which for these first three categories was receiving more than 365 test strips in a year, which is testing more than once a day and then testing more than four times a day for the short acting insulin group. Well, we had a number of sensitivity analyses using more conservative threshold, but when we identified that potential excessive use, this is what we find.

Here it actually plotted the percent of all test strips that are dispensed in excess. There were a total of 157 million test strips dispensed between VA and Medicare. You can see there is no diabetes medication group. Among individuals who got their test strips from VA only, around 15 percent of test strips were in excess compared to those who got their test strips from both VA and Medicare, almost 50 percent of all test strips were in excess. These findings were adjusted for diabetes severity, which we used for both Medicare and VA data.

Really, the findings are less important here than just the aspect of identifying medication use, comorbidity test strips, etc., from both VA and Medicare simultaneously, to look at healthcare use among veterans. That is it for the examples. A few slides on where to go for more help. Then we can get to the questions. They are great, great data sources. They are great sources of information. Some of them are listed here. VIReC maintains a website specifically on VA CMS data, which is extremely helpful for anyone interested in using these data. VIReC also has a web page focused on pharmacy data. There is a very nice user guide. It is a little bit old at this point, but it is still very helpful for understanding the different kinds of pharmacy data. What is not listed here but is actually also very helpful is the Vinci website and the information about CDW. There is the List Service, which many of you are probably familiar with. Then the Help Desk. Although I have not used the phone number. I have used the email a lot. It is very, very helpful. I highly suggest soliciting VIReC’s help doing this kind of work. I think we are just at 40 minutes, which is great. That is where we are going to end. We have a lot of time for questions about any of these data sources that we are happy to answer. I will turn it over at this point.

Moderator: Great, thank you so much. We have a few questions here for you. I will just get started. The first one: is there a way to determine whether a drug is generic versus brand name via analysis of the NDC code?

Walid Gellad M.D.: Okay, yes, there definitely is. That is the reason why we used NDC codes for our study in VA data. We had a study a couple of years ago where we compared brand name use between VA and Medicare. For some drugs within VA, it is quite easy to identify brand and generic just on the name, because there is no generic form available. In some cases, if for example the drug is like Enalapril and the VA product name will say like Enalapril, but you cannot be 100 percent sure whether it is the brand or generic. In that setting, NDC is actually very useful because the NDC code will tell you whether the drug is brand or generic, not with certainty, but there are algorithms for defining what is a generic and what is a brand. The answer is yes with NDC codes.

Moderator: Has the pharmacy data been combined across sources? For example, drawing opioid use data from VCMA files, CISR files, PBM files?

Walid Gellad M.D.: I do not know the answer to that. I think there is that report that we showed comparing drug use between reported prescriptions between DSS and PBM data. I have not seen a study specifically of opioids, but we have inpatient and outpatient data using CDW. I guess that is what I would say. All of these data sources, like Xinhua mentioned, get their data from the VistA files, the local VistA file at each VA. Presumably, you are picking up the same prescription, whatever data source you use for the outpatient data no matter whether they use PBM, DSS, or CDW.

Xinhua Xhao, Ph.D.: I want to add a few words here. VCMA is inpatient pharmacy data. It is not outpatient pharmacy data. You will want to make sure you only use inpatient or outpatient or both because all the national pharmacy data sources, they extract the raw data from this part. We should expect high consistency between those data sources. As long as the variables you need are available from the three data sources, like which you actually use should not matter very much.

Walid Gellad M.D.: Let me just add one thing. We had mentioned the issue of sig not being available in DSS is important. If you anticipate a need to know what the actual directions were on the prescription, that would be PBM and CDW data would be most helpful. There are some issues with identifying the number of dispensed doses in DSS data because it does not tell you what the dispensing unit is in the DSS data. You need that sometimes when you really want to figure out how much drug was dispensed.

Moderator: Okay, thank you. The questions are really coming in so I will just keep going. Has VA pharmacy data been used for randomized control trials?

Walid Gellad M.D.: I do not know the answer to that. I would love to know the answer. I assume you mean for outcome assessment. I do not know the answer to that. I think if they have not, they probably should be. The issue is, again, the timing of the data, which for VA data, you can get fairly real time. If you want to go and understand non-VA data, there is that data lag, which is the problem.

Moderator: Okay, is it also possible to get diverse drug reactions data along with the pharmacy data?

Walid Gellad M.D.: I do not know the answer to that. Xinhua, do you know?

Xinhua Xhao, Ph.D.: For what?

Walid Gellad M.D.: Adverse drug event data? I think it is not in the PBM data and it is not in DSS. It would be, if anywhere, in the CDW data, but I do not think. The VA does maintain an adverse drug event reporting system, which is maintained by the PBM, but I do not think that they offer that the researchers. I am not 100 percent sure.

Moderator: Okay, do any of the sources contain the structured and non-VA medications?

Walid Gellad M.D.: There is a set of raw data tables in CDW that includes non-VA data. It is going to be raw. There is not going to be any kind of. It is just whatever is typed in and present in the VistA file. That is what you are going to get in the CDW data table. There is a specific data table that you can request of raw data that includes what someone has entered as a non-VA med in their local VistA file.

Moderator: Are there any publications on the percent of VA patients who obtain their medication outside the VA?

Walid Gellad M.D.: There are some that use local data or surveys. There are some for particular drug classes. Then there is the large VA survey every year that they do what percentage receive their prescriptions from outside the VA. The answer is yes. Now whether it is national and recent, I do not know what the best source is to point you to for a recent national number of the percent of VA patients who obtain meds outside the VA. I am sure it exists.

Moderator: We still have quite a few questions to go so I will just keep going. Where do we find the drug code identifier MediSource?

Walid Gellad M.D.: Can you say that again?

Moderator: MediSource, M-E-D-I Source.

Walid Gellad M.D.: MediSpan?

Moderator: Source. I wonder if it is MediSpan.

Walid Gellad M.D.: Well, I will just answer it assuming it is MediSpan. MediSpan is what we use. There are others called First Data Bank. Some people use the Red Book. You have to pay for these databases. You have to pay for it. It is a license that you purchase. Then you have access to it for a number. You can look up. In MediSpan there are these therapeutic categories that they create so you can look up, for example, diabetes drugs. It will tell you all of the diabetes drugs and their associated NDC codes. That is what we have used MediSpan for. I am happy to help anyone find it. You just have to purchase it.

Moderator: Is there a way of assessing issues of medication compliance from these data sources?

Walid Gellad M.D.: My mouth is tired. Do you want to answer?

Xinhua Xhao, Ph.D.: Substance use?

Walid Gellad M.D.: No medication compliance, adherence. Yes, we have done medication adherence. Yes, we have done medication adherence. Other people have done medication adherence. You would want to use PBM data or CDW data where you can find days supply. You can find the quantity dispensed. You can also have the sig in case there are questions. Then you use the usual calculations that are done from claims data about medication possession ratio, or proportion of days covered to identify adherence, which Xinhua has definitely done. It involves sometimes complex programming and accounting for prescriptions that overlap across the year, but the answer is yes.

Moderator: How is the Medicare Part D Slim File data linked to the VA pharmacy data?

Xinhua Xhao, Ph.D.: Okay, let me.

Walid Gellad M.D.: Go ahead.

Xinhua Xhao, Ph.D.: Okay, to link to the Medicare Part D data and the VA data and we need to link by the patient level and service state level. You can link the two data fields.

Walid Gellad M.D.: You have an individual who is present in both Part D and VA and then you can look at the dates of sales for examples, but they are linked based on patients and VIReC is the one that will do the linking for you when you request the Part D data. They will send you the Part D file that is relevant. Then you link them.

Xinhua Xhao, Ph.D.: The procedure will be you create a dummy ID for each patient for your study cohort. Then the dummy ID and send that to VIReC and they will send back data with the dummy ID variable.

Walid Gellad M.D.: Let me, because I think it is actually worth talking about what the Slim File is. We did not have a slide in there for that, but the Slim File is a select number of variables from the larger PDE file that Medicare, CMS has. You can get all of the important variables in the Slim File including the brand and generic name, the date of fill, the quantity dispensed, etc. What you do not get are things like whether the drug was subject to quantity limit or whether the drug was subject to step therapy or prior authorization. What phase of the benefit design the patient got the drug. Those are things that are not in the Slim File. There are also other Medicare Part D files including prescriber characteristics, information about the pharmacy, information about the formulary that is available, but is not available through VIReC in the same way that you get other Medicare data. You have to do a separate request process. There is an extensive cost associated with it. That data is available, but those are the issues. It takes a long time, six to 12 months. It costs a lot of money. Whereas the Slim File is much more streamlined, but you only get those 14 variables.

Moderator: Okay, many researchers have wanted to connect on medication with the provider type or the clinic associated with it. These fields are either poorly populated or simply say pharmacy. Do you have any idea how you can link a prescription with the prescriber?

Walid Gellad M.D.: That is a question we have been wrestling with a long time. It is clearly feasible. PBM has the prescriber information. I will just focus on PBM data. PBM has the prescriber information for every prescription. They use that internally. I think that technically, at least in the user guide it is available although we have never really gotten the prescriber identifier. I guess it is kind of a vague answer. Technically, it should be possible. I think with the right approval, you should be able to receive that data. That is in PBM. In CDW, it is the same issue. There are fields for who the prescriber is, but to access the actual scrambled social of that prescriber requires authorization. You do not want to just use the identifier that is in the file because identifiers can be used in multiple VA’s and it can be recycled. They are just not ideal. You can always tie a prescription to a concurrent clinic visit and you can know the stock code for that clinic visit. The PBM files we use, I think we did have whether it was a nurse practitioner or resident or an attending. We have that in the files that we used a couple of years ago. We have not yet had success in the general way of doing things of just being easily able to identify the prescriber. Let me just add also in the Part D data, Part D is now identifying prescribers, but that is not something that is going to be in the Slim File.

Moderator: All right, moving on, what is the best source for IV chemotherapy data, which is typically given in an outpatient setting?

Walid Gellad M.D.: I do not know the answer to that. I am sure it should be PBM will have IV medications that are dispensed. I do not know if that is the best way to look at it is PPT code for an injection or whether it is to use PBM data and look at outpatient prescriptions or whether it is CDW tables. I do not know the answer to that. I would say the best person to ask is probably VIReC.

Moderator: Which of the VA data sets include prov code that we can access? The person says that my son says no for PBM, yes for DSS, NDE, NCDW. This could be an important difference in the data sets.

Walid Gellad M.D.: What was the question again?

Moderator: Which of the VA data sets includes a prov code, P-R-O-V code that we can access?

Walid Gellad M.D.: Is that the same issue about providers?

Moderator: It says P-R-O-V code. We can skip this one.

Walid Gellad M.D.: Yes, if someone wants to rewrite the question, but I assume it might be about providers. Let us skip.

Moderator: Do any sources contain information about whether medication listed has been discontinued versus just expiring?

Walid Gellad M.D.: If anything, that would be in CDW. I do not know.

Xinhua Xhao, Ph.D.: You meant the medication is continued?

Walid Gellad M.D.: Yes, the DSS and PBM, those are dispensed prescriptions. Anything you get in those files will have been dispensed. You do not know when it ends, whether it ended because the patient decided or the provider decided or whatever. You do not have an indicator of it being discontinued. I do not have enough experience specifically CDW to tell you if that is something you can find in one of the tables.

Xinhua Xhao, Ph.D.: I want to add one point here. The PBM data set, we have variables to indicate if it is a resale or a new prescription, but to identify if it was discontinued or not, we adjust the program and it will check the discontinuance of all the prescriptions based on the different technician. The cost part, you may want to use a certainties gap, the difference in the cost point to define discontinuation. That is one way. The variable exists to tell you the drug and whether it was discontinued or not. You may need to program to find the discontinuation.

Moderator: Okay, I am going back to the last question. That question was about provider IDs. The person just sent the comment, “My understanding is that PBM is not authorized to release this info, but DSS and CDW is.”

Walid Gellad M.D.: That is very helpful. I think that reflects our bias that we have typically used PBM data. I think as long as you have the right authorizations, than you can identify who those prescribers are in the CDW data and it sounds like the DSS data, too. That is very helpful to hear.

Moderator: How much does the Medicare Part D Slim File cost and how much time does it take compared to the larger Medicare Part D data sets?

Walid Gellad M.D.: The Slim File costs approximately zero dollars. That is what is nice about the Slim File. It is like the other Medicare files available to VA researchers for no cost. If you are talking about the larger file, not the Slim Files, that you request in that special process, you are talking about 20,000 dollars depending on the size of the data, 15,000 to 20,000. At least that was how it was five years ago when I last looked.

Moderator: Okay, I am going through the questions to see if there is anything else. There is one more question. When a new prescription is given on a date that was before the previous prescription ended, as shown by Dave\_Supply, do you assume the next one is taken after the previous end or the previous one is discontinued?

Walid Gellad M.D.: That is a great question and I think that is done differently in different studies. You really have to look closely at the methods. The release date, which is what we typically use. There are two variables. One is release date. The other is fill date. When we look at release date, if it is a window prescription, it is the date the patient picked up the prescription. If it is a mail prescription, the release date is the date it was mailed to someone. Technically, the time when the pharmacist scans the prescriptions and it is dispensed. It is mailed, there are going to be a couple of days before it actually reaches the patient. I think some people will just look at drugs “as prescribed” when it was released, that is when they got it. Other people will say they will give a two to three day grace period and assume that it starts two or three days after the prescription is mailed or they will say if it is the same drug and the same dose, we are just going to start it when the days supply of the prior prescription is actually run out. There are lots of different ways to do it and people have done it depending on the specific question.

Moderator: Okay, we are just out of time. Heidi, do you want to take things over?

Heidi: Yes, we are going to close things out here. To our presenters, I really want to thank both of you for taking the time to put together and present for today’s session. We really appreciate it. For the audience, I want to thank everyone for joining us for today’s cyber seminar. I am going to close the session out in just a moment. When I do, you will be prompted for a feedback form. If you could take a few moments to fill that out, we would appreciate it. We really do read through all of your feedback. I want to thank everyone for joining us for today’s HSRND cyber seminar. We look forward to seeing you at a future session. Thank you.

[End of audio]