>> So, we're going to look at now how to use a PDE file to examine medication utilization which is often of, you know, the topic of interest when you do some prescription drug research and here, in this section, I'm going to look at only--discuss two things. The first one is what variables in PDE data are related to medication utilization. This is basically actually repeating what actually Marshall has covered in the previous section. And then the second topic is we're going to actually look at the measures of medication utilization we can construct based on PDE record. It's more about the--not--It's not really about the data, but it's more of a research topic and we are going to discuss them.

OK, so what variables in the PDE are related to medication utilization? Well, first, if you want to research medication utilization, you'll have to identify why these are the drugs that I'm interested in. So, in the PDE file, the Product Service ID which is essentially NDC, NDC level, it gives you information about, you know, the dispense the drugs. So that's the first one. And then, it's the [inaudible] of this variable which is a branded name and generic names. But if you want to use this brand name or generic names, again, you have to spell correctly in the--in a way--in the same way as the PDE data actually spells these names. So these are the potential variables you could use to identify the record or observations in the PDE data that actually fill this--the drugs that you are interested in. OK, that's fine. If you have only one single drug or pretty much a couple of drugs, then you could use this outfit, but what if again, what if you are interested in examining therapeutic class? So again, it goes back to Marshall's point where for this case, you might actually need some commercial database which tells you about the--which actually already identified in this code for a specific therapeutic class. So Medi-Span is the one example. Well, another case is like again, Marshall said, they're like ARB or statins when there are only a couple or several generic names or--that belong to that specific therapeutic class, then you could potentially use its generic drugs. Well, statins, I once identified, I think that they're a--it's a Furadantin [phonetic] including combinations. So if you know those generic names, then you could use well these other statins. You could use all that variables.

OK, now, there so--Now, you know, what drugs you want to examine, you are going to look at and then to look at the medication utilization and what informations do you need. What do you want to know? When the prescription was filled and how much? So the when is in the variable code of prescription service date. So this variable tells you the prescription fill date. Actually, in your note and handouts, there is no fill there. There is only probably initiation, but this--precisely, this is not initiation date because we don't know from the PDE data or--and you probably claim data there is no way that we can actually identify the initiation date. So this is, well, basically prescription fill date. It's a service date variable. And then the second variable, the prescription data supply variable. So it gives you information of how many days of prescriptions were filled. Again, this is filled, not actually used. So this is about our prescription amount.

So this is a quick descriptive data about this data supply variable, so

which is a key information if we want to construct any medication utilization or a medication adherence measures. So the median or mode of this variable was at 30 days, which makes sense because about 70 percent of PDE records are tablets. So basically, this means that the beneficiaries fill their prescriptions on a monthly basis. That's a very common way. And then about--Now, the next about 20 percent of people fill the prescription between 1 to 29 days and then another 13 percent, about 13 percent had more than 31 days. Sometimes they use mail order system then probably it's going to be three months. Use it three months is pretty much common. And--but--and then a small number of prescriptions they fill are actually greater than 100 days and--but when you work on this variable, you're going to find the records with variable zero. So what this mean, that prescription is the supply days is at zero. Or sometimes you may find a 360 days or 200 days. So, these are very, very small number of observations. It was less than 0.1 percent of the--or the PDE records. So these could be actually typo. It's hard to think of, OK, supply days, you know, days of supply is zero. And when I looked at all this zero record and I actually found out that they are mostly--not all of them, they are mostly those who are in employer's master plans. So if you remember when Marshall covered the--this plan type, so they are not required to submit PDE data. So you may see some PDE records, but then they are not. So they may just to plug in zero or I don't know, they didn't pay attention or [laughs] so, yeah. Right, so it's hard to utilize this as zero variable.

But on the other hand--Or these are--What about then these 300 days or 300--no, 200 days? They could have simply type or they could have simply outlier, but depending on the drug you are--you are examining, it may not be. So don't discard right away because let's just suppose the drug is cream, it's not tablet. Cream, it could be OK. You could have used this cream for six months or for a month. So it just depends on your research or your study or the drug. So look at these records in case in quantities.

All right, so now, the next topic, we're going to move to measures of a medication utilization. And really, the basic way of actually examine or construct this medication utilization is where as Marshall said or you discussed, we--you already identified your cohort. OK, my cohort is patient with diabetes. Then, now, probably you link this patient, these records with the PDE data and look at what--whether they are taking or they are like fill the prescription [laughs] with antidiabetics, right? So in this case, this measure is simply to find an indicator whether a, you know, represent whether the patient filled the specific medication you're looking at and this can be used for a population with a specific condition, right?

So then what about in PDE data context or Medicare context, you need Part A and B data in order to do this which--meaning that you can't do this for MA-PD enrollees. So Part A and B data are available only for those who are in the tradition of people services sector and who are in standalone PDE--PDP brands. And so, this is a little bit restricted. So this measure actually really requires information on diagnosis. And then, another way to examine medication utilization is to construct a medication adherence which captures whether a patient that takes a prescribed medication according to schedule. It's to say where a doctor said that, "Well, take

this medication for a year." These are medications for chronic conditions like statins or antihypertensive. OK, then we want to look at whether they are following the schedule or not. So for this--to capture this medication adherence, various measures or several measures have been used and why several measures and how are they different. So basically, all the measures actually capture the basic concept that where--are they following the schedule, but on the other hand, it depends. They differ depending on whether they are focusing on the percent of time with the prescription or percent of time without prescription, but it's your choice. You know, more interest. It does--a different way of saying it where these patients stick to that medication for 80 percent over time, are over one year. And then another way to say is where 20 percent of time they were out of the drugs, right. And also, these measures can be defined differently depending on how you specify your time period, and we are going to look at how this can be. How much difference this actually, you know, this time period, how to define time period can make a difference. Right.

So, we are going to actually look at this medication adherence, but before that, I'm going to actually put--just to put things into the Part D or PDE data context. Again, I'm just reminding you these three key variables, you need to use to construct any measures that I'm going to discuss. That's the three variables, the Product Service ID and service date variable, and days supply number. The question is whether this is imputed or computed of the text--have computed or imputed the variables. So no, this is just simply, it tells you how many days of drugs you actually filled.

So, again, some limitations actually using our PDE data to construct any adherence measures.
^M00:10:06 So this--Any adherence medication, adherence measures that we construct based on PDE data, they capture acquisition or possession of a medication. They do not actually capture consumption. It's just a repeating. And the second, we also can't measure the timing of taking medications. So some medications are sensitive to timing of taking medications like the proton-pump inhibitor, PPIs. Usually, they're recommended to take early in the morning before you eat anything, or antihypertensive, usually recommended to take at night, but there is--we can't actually get timing information from PDE data. So for these information, when you're really interested in just the consumption itself, then usually it isn't from--if you look at the studies, they have come from surveys. So, but then--but the surveys are usually expensive and it just covers only a short time period and you have relatively small sample size. So, but on the other hand, using PDE data file has all the way instead, it allows you to have a large number, a larger sample size and also you can do long-term follow-up study. But again, these are the limitations of using our PDE data.

OK, here is the summary table of the measures of a medication adherence. This is actually really a summary. There are more than that probably. And these measures, actually, I focused on measures so that actually captures some medication availability. In other words, measures that capture the percent over time with medications, right? And--but that basically these measures, all these measures capture the proportion of days of supply during

a specified time period or the time period between refills. So we are going to get to this, how we can actually capture this. And there are actually several measures can be defined or constructed to capture medication adherence. The most common measure is MPR, the Medication Possession Ratio and PDC, Proportion of Days Covered. So, and other measures--These measures are--It's similar to MPR, but they are slightly different probably depending on the study or data settings. So depending on the--your study specific or your [inaudible] interest, your research interest, you can--you may just slightly define your MPR, but basic idea is we want to capture the percent over time with our medications. So if you want to know the definition of each--specific definition of each measure, please look up these two studies. These are both our review studies and they actually provided specific definition of each measure and also they provide measures that focus on medication--that percent over time without medication. They just simply selected these measures to focus on MPR and PDC.

So MPR and PDC, I'm going to come back to these during the exercise. There are slight differences in definition between these two measures, but on the other hand, my impression is it's a study specific. You can divide this, but try to be clarify how you measure your medication adherence. And some people, there are some conventions of defining these two measures, but on the other end, that some people actually follow the definition of a PDC and then--but they call them MPR. It's a bit sometime a little bit fuzzy area so, but just to be sure that you actually clarify. OK, this is the definition or before that, make sure they're what you want based on your research or question.

OK, so then let's look at MPR. So MPR is one of the most commonly used measure of medication adherence and it captures the proportion of days of supply during a given time period and it can be defined differently. Now, it's depending on how you specify a time period. OK, here are the two main definitions or two [inaudible] definitions. So the first one is MPR, it's just a conventional definition. If you see here, it's simply total prescription days of supply during a specified period divided by total number of days during this specified period. Well, wait a minute, how do we specify period? It's basically your study period, right? It's just that during my study period were, that the total days, total number of days of my study period was something, but then out of that, how many days the patient period of prescription? So, there are two ways to specify your time period or define your study period. So I'll come back to the MPRm later.

Let's look at how we can specify time period. So, the first way, uses the--have the fix at the follow-up period. So everyone in my data set, I'm going to follow them up, each patient for one year. So one person, one beneficiary began to take statins January 1st on 2013. OK, I'm going to follow up this beneficiary until December 31st on 2013, fine. And then another person began to take statin February 1st. Now, you need to follow up until when? January 31st of 2014. All right, so everyone has exactly the same fixed time period. And you'll say that we followed up our patient for one year, right? And then in this case, so actually, the interesting is that the numerator is [inaudible] which is basically duration of a

therapy. It can be used as one medication adherence, well, measure of medication adherence. For example, as [inaudible] as she said, OK so out of 103--365 days, there was only two prescription periods, 30 days, 30 days, now to 60 days. Then 60 days divided by 365 days. Or you can say the work duration of a therapy for this patient was 60 days and another patient that probably have 180 days or 300 days. So, that's one way you can do that. OK, another ways that now as in even in our example, this--actually, the first fill date is different between this patient, right? So in order to follow this one in the--If you want to buy or purchase this PDE data, you need two years of data sets. What if some patient began to take statin in September? So you need to follow up until August, right? But then there are some time, you know, data constraints or if we are doing some clinical trial, sometimes you have to close all your study at some point. I'm going to recruit my patient just until this December 2013 period. So when you have to close your study or when you have only one year of data, sometimes you are limited and you cannot really follow up exactly same dates for each patient. So for this case, you can actually simply say that where for study period, it differs across patients. So we're going to follow up each patient from the first fill date to the end of a follow-up date. To the end of a follow-up date period means they have to--until the end of--to the end of my study because I'm going to close my study December 2013, right? So if this--In this case, the denominator can be very off by patient. Is it clear? Yeah, right. OK, so that's MPR.

So now, if you know that you--how you specified your time period and then use it to get the number of--total number of days which could be--vary by patient. And then in the numerator, you simply sum up all the days of supply during that period, right? Then let's look at this MPRm which is--which I named MPR modified. So look at this denominator. So now, this denominator says that total number of days between the first and the last fill date. So, let's continue our example that the patient took statin first on January 1st and then the last time that she showed up in my PDE data is August, right? Then what should I do? So in our previous example, in MPR case, it's--where out of 365 days and probably seven months you took [inaudible] prescription. But here where I'm going to define my denominator as the duration between these refills. So from January 1st to let's say August 1st, that was the last time she filled the prescription. So during this, then basically seven month and how many days supplied? I mean that the patient filled the prescription, and probably even between this January to August, it's not necessary this patient actually filled the prescription every month, right? She may have just filled the prescription, let's say January, 30 days and then come back on April and filled the 30 days and come back in August and then filled the 30 days. So in between this, January to August 1st, actually the total supply days is 60 days only because you can't really count on the last one because this is interval, right? So really was that this can be--sometimes this can make--shows a difference. And how does it really meet this--the denominator make a difference.

^M00:20:05 In other words, how we define our time period or our denominator, does it matter? So this is summary of the similar example that I just described here. So this study period is actually one year. So let's say

I'm going to, you know, recruit my patient starting from April 19th and then close to 2009 and then close April 18, 2010. OK, so this beneficiary first filled her prescription on April 27th. And in the last day, she filled the prescription is August 25th. So, let's try that the first definition, our conventional imperative definition. So in this case, your study period is one year, right, 365 days. So where out of that one year, 365 days, you can sum up all these days of supply values to total prescription days of supply, then 118. So, 118 divided by 365 days. That's the usual MPR. So that's the most common actual measure, the way to measure a medication adherence. So that MPR was 0.3. Or if you want to say, "Well, I'm going to follow up this person from the first fill date to the end of a study date," probably you have to subtract 9 days out of your denominator, which is going to be close to 0.3. It's about 0.3 or 0.35. Right.

So that's MPR, but what about MPRm? What would be your denominator here for MPRm, MPR modified? You want to focus on this duration between refills. So that was basically--the definition was the total number of days between the first fill date and last fill date. So, May, June, July, four months. So about 120 days. And then, so you can divide it by this 100--no, that it's now 118, you have to subtract this 10 out of this one. So about 108 divided by 120. Wow, MPRm is 0.9, that this patient is pretty much compliant to his medication, right? So what makes this difference and how do we choose? It's a tough question, right? So, which one would you choose and why? Well, the denominator makes the difference for sure, but what really the core fundamental difference in measuring different way? Any thoughts? Mary, probably you have constructed this?

^M00:22:59
[Inaudible Remark]

^M00:23:00
Oh, she said the chronic [inaudible]--Oh, OK, so the fundamental difference here is that how we want to treat early termination. So this study period is basically it's one year, right? We wanted to follow up this patient for a year, then when we actually calculated the MPR, we treat it, this early termination as discontinuation or a nonadherence or a noncompliance, right? So, she said that chronic condition, yeah, it maybe make sense for statins because when some patient is--or on statins, they usually--it's likely that they are recommended to be on the statin for long time or for the rest of their life. It's pretty much long time. It's not like meant to be three months or six months shortened therapy, right? Then, the second case that MPRm, MPRm, the modified measure is, well, how do you know whether this is a really nonadherence or a discontinuation. But this maybe, she may have some adverse effect. And physicians have to restart her [inaudible] or stop taking medications. That's possible, OK? Or she may have some allergic reaction to medications. Then in their case, why they're made to this [inaudible] that we calculate may not really present medication adherence. So it will again, it goes back to what drugs we are looking at or what your research question is, right? OK. Right.

So, how do we choose the units measures and in [inaudible] paper and he

said, where how to treat early determination. This is really key and it depends on whether therapy is generally prescribed for a short time period only or it's meant to be a lifelong condition, to treat a lifelong condition. So then we can actually choose for the--whether we choose MPR or MPRm. We will actually look at the real--In the exercise, we are going to look at PDE data and compare this MPR and MPRm, right.

OK, so another a little bit sort of issue here is that let us suppose you decide to use MPR and you are comparing two groups and in one group on average the MPR calculated was 0.8 or 80 percent. It seems pretty good. And another group had 0.82, 82 percent, are they really different? Well, besides from statistical significance, which is often function of your sample size, so you do want to really say that one group had 80 percent of adherence rate, another group had 82 percent of adherence rate. It's hard to say, right? So that in other words, the question is does it really make any clinical difference between 80 percent to--an 80 percent and 82 percent? So then, so one way--So people have done, taken some dichotomous approach, in other words, it is embracing there, where if your MPR is greater than 0.8, we are going to define it as good adherence or this patient is compliant with the, you know, medication. But it doesn't matter whether it's 80 or a 90, an 85, yeah, maybe, just to say that they're pretty good. But if it's below, it's less than 80 percent, yeah, maybe they have a poor adherence, right. So, that it has been a really common approach actually taken in the literature, but in that case, probably you want to actually justify a little bit of your cutoff point.

So for statins' case, I have seen many studies that you will use the 0.8 as a cutoff point based on the evidence that actually made some--just whether your adherence rate is above 80 percent or below 80 percent, it's indirect that they have a little bit differences in clinical outcomes. So besides that, if you have other drugs, then you may want to look for some evidence or literacy and justify your cutoff point.

OK, and now, another is actually it's a really practical issue. It's simply data availability and cost. So why MPR is popular and why people use MPR or related measures? Any measure that I showed in the table was relatively easy to calculate. You looked at the definition, yeah, use this to sum up the total of supply days and you know the denominator and you can get it. And also, relatively easy to interpret, but I just put it in the parenthesis just to interpret because there are issues related to medication, the MPR measures.

So something then you want to keep in mind that when you--what you are doing, OK, well, what kind of restrictions or limitations this measure have, right? So we are going to look at those. The first issue that you need to consider when constructing these measures is oversupply. What do I mean by oversupply? Any idea about this? What do I mean by oversupply? Or some people say, overlapping supply, the same, or excessive supply. Yeah, I'm going to move to actually the next slide. I will come back to other points. Here is the example of excessive supply, supply or oversupply. So let's look at the data by the--the line by line a little bit. So the first date that this patient fills medication was at June 19 and days of supply was

35 days. Well, then when did she come back? July 10th, so what's the days between these refills? Yeah, less than 35 days. Yeah, it's only 21 days, right? So, maybe--well, it is a possibility that this patient actually lost some of her pills, I mean, who knows, right? But--Or maybe she's holding on something and who knows, right? But anyway, so that's a possibility. Maybe she's a little--so let us suppose that she is actually carrying these 14 pills around with her, and then OK, so this one looks good until we come to--around here. Is it--December 18th. So she filled another month of prescription drugs for--in--on December 18th, but she came back on January 2nd, right? So when you sum up all these data supply for your numerator, basically, for your numerator and it's 400 days. But study period is 365 days, so what's your MPR? It is greater than--
>> Greater than one.

^M00:30:05
>> Greater than one. What does that mean? So how are you going to interpret? So, of course, when you use dichotomous approach, no problem. Yeah, it's great in 80 percent period, right? It's relatively easy if you use dichotomous. Or if you--Most of the time, actually, in a really quick fix which you are going to do in your exercise, this is not just a truncate. I mean, we really don't know what's going on with this sort of oversupply, so--but this is issue sometimes. And you can actually look at in the PDE, their 2008 PDE data statin users and appear sometimes greater than two or three which is very interesting and you will see it was, OK, this is in your data in the exercise. And the difference between MPR and PDC actually comes in actually how we want to address this oversupply. It's a very subtle difference, but we'll look at that.

OK, then another--The second case is what if patients were admitted to hospital? Hospitalization, OK. So we are going to--How do we want to do that? And then another issue is, well, what if poor conditions that require multiple medications during specified period? Then what do we mean by medication adherence? I mean, how are we going to address this? So let's look at one by one here. OK, handling hospitalization or SNF, skilled nursing facility. So any time--Either way, so when patient who were sort of institutionalized, how to handle this? So when a patient who was hospitalized, it's very likely to me that they were prescribed. I mean, they were taking the medications, right? It is very likely, right? But the problems that they don't show up in your outpatient claims or this PDE data in the Medicare, right, because it's an outpatient based, right? So we don't--we really can't capture those things. And then, this patient was hospitalized, but then you are summing up over this, your total data supply just based on PDE, then you are likely to underestimate the true medication adherence, right? Then what do we do?

^M00:32:22
[Pause]
^M00:32:25
Or simply--I will use the--
^M00:32:28
[Inaudible Remarks]
^M00:32:29

Yes, you are--If you want, yeah, yeah. Or even--first of all, whether we--we need to identify whether this patient was hospitalized or not, but then you need a Part A data. When you don't have, then it becomes a limitation. We are missing sort of a hospitalization. In Part A data, is there any drug information?

^M00:32:51
[Inaudible Remark]
^M00:32:52
No, they--No, they do not have any information. OK, so then, let us say anyway that as she asked, that can we know that--what kind of medication that patient took when they were hospitalized? No. And can we know that--whether they prescribed extra days of medication when they're discharged? No, nobody can do. So, let us say we have Part A data and at least we know that this patient was hospitalized and then the days of hospital stay, then let's assume that they took this medication where they were at the hospital and we are going to add that number of hospital stays to the numerator. Or if you know that, I'm going to reduce my denominator by this number of hospital stays. That's a--either way, right? So that's one quick sort of a fix for this hospitalization. Here is the example here. So 28, so this study is again following up one year, all the patient and then one patient had 280 days of prescriptions, but she had 20 days of hospital stay. So if we ignore this one and, well, I don't know, we don't have information on hospitalization, then you calculate the MPR, 280 divided by 365 is 0.76. In this case, well, I'm going to add these 20 days to numerator and that this is--well, I'm going to reduce my denominator by this number of hospital stays and it gives us a little bit similar, 1.82 or 0.81. So if--when we ignore it, well, if you are using continuous MPR, continuous measure of MPR, 0.76 versus 0.8, may or may not be. But if we are using dichotomous measure, it could make a huge difference. This patient is now considered as a poor adherence. And I know if you account for this one, then it's now good adherence, right? It could make some difference here.

All right, so the last issue. So, what about multiple medications for a given condition, how would you handle?

^M00:35:10
[Pause]
^M00:35:15
So, one possibility is, if we were really interested in, say the conclusion, overall adherence is blah, blah, blah, then you need to come up with some measure, some composite or one measure, right? So, in this case, well, let's first create prescription or medication-specific adherence. So for medication A, you calculate the MPR and for medication B, you calculate MPR. And then, yeah, it's somewhat crude, but get the average, on the way to the average. Well, let's say, it's the average. What's the benefit of this approach? Simple. This is simple, right? The quick way of actually coming over to some composite measure of the medication adherence and it summarize actually overall adherence, but what's the cost? If it's a one medication is--That the question is, well, what if one medication is really important and the other one is not, and what--then if that's the

case, you could have probably used weight. You can create that. That the real problem is what if the adherence is really different between these two medications? So maybe, in underlying issues, maybe one is important or maybe the other one is not important, but then you're right for they are really different. For example, for medication A, the difference is 0.6 and medication B, there is--MPR is, let's just say 100. What's your average? 0.8, yeah. On average, you're saying that, well, this patient has a good adherence to the medications for this condition, but no, you're actually misrepresenting adherence for one drug. It has poor adherence for medication A, right? So in case that whether it's because a better one is important and the other is not, when you expect some huge differences or when you actually calculate this medication-specific adherence and you see it as the average. Just a summary statistics where one is a 0.3 and another is a 0.8, then average that this supports may not be really appropriate, right.

So then, well, if you don't want to get the average, then what do we do? Yup, let's just say, let's not make any, you know, conclusion comment on this over adherence. We are just to say, well, for medication A, this was adherence and for medication B, and this is the med adherence, right? So, you--again, you create medication-specific adherence and they model separately. But you are looking at sort of race differences or age differences where for medication A, there is no difference while for medication B, where there is difference. That's another way of doing it. But then, while this allows actual adherence to each medication to be represented which is nice, but on the other hand, you may need a little bit complicating modeling because the error terms--of course, now, you have multiple equations to at least the two equations or three equations, then the error terms of--in each equation, the--of course, the equations are likely to be correlated. So, when you don't account for this correlation across error terms, then you may inflate or overestimate or underestimate a variance over your--these estimates which, again could be misrepresenting, right? So, it's--it requires a little bit of modeling.

Oh, OK, there was another one. OK, timing. So, so far, I have sort of increased the--assume that you are going to use a person-year as your unit of measure, but why? You could use a person-month, right? So it's another decision that you have to make where do I want to look at the medication adherence by year or by month. Person-year, yeah, it's again simple, it's really simple. You have 2008 or a 2013 PDE data and get the medication adherence measure. The MP--calculate the MPR for this year, then, yeah. And let's say the word, that's a really easy way to do it. But on the other hand, what if some medications may have really variations across a month. So when there is within year variations, then you are missing some information. And in Part D data case, for example, when people are likely to stop taking their medications? When they fall into the doughnut hole. So when they hit the coverage gap, maybe the medication adherence could sharply drop, right? So this hitting the gap or it depends on patient, but it's likely to occur towards the end of the year.

^M00:40:05 So maybe you are expecting poor adherence in October, November and December, then you may want to calculate this person-month adherence,

right? So in this case, you can actually capture that--to some variations within a year, right? But, again, this requires a little bit more complicated modeling because now you have 12 observations per patient. Again, this is a repeated measures, right? So for one person, January to February and then you need to control these error terms.