

>> Beth Virnig: Are you all excited about [inaudible] like my boss told me I had to come. I don't know why I'm here. And so we started off doing this thing, it's like, "Well, why are we here, and doing some really big picture reality check about what it is you can and cannot do with these data?" Again, big picture. Before we get in to the details of -- this is the file, here are the variables, here are and how they're used is to really make sure that we're thinking conceptually both about -- and you noticed, we start with strengths because it's really about where are the opportunities and then where are the places we might want to hold up and be a little bit more cautious, because these data are really hard. And, you know, they're challenging because you can always get something. Okay? It's like any big data file. You can run a regression and get something and get a P-value. And what I would personally say is there's all those technical problems, the hardest part I think is thinking about the data and understanding what it is you have and making sure that what it is you actually found makes sense and is something you want to be standing up and taking credit for.

So the origin of the data, it's derived from the reimbursement of the payment of bills, and so that's really interesting and at very, very early days of Medicare data, this was like the fatal flaw of the data is what they came from, sort of fruit of the tainted vine. Like people will say anything to get paid and therefore you might as well just do regressions and random numbers because that's basically what you have. I think the literature has dramatically moved since then, but if you go back and you look at any articles, if you ever do that like a completely thorough lit search and you look at articles from the early '90s, you will see these big long apologies about the fact that these were administrative data and that people will say anything to get paid, and they might not have done it, and it might not even be the right person. And I think there have been enough validation studies since then that people have really backed off of that. But in fact, the fact that it's from reimbursement bills actually provides some advantages to us now. But -- so we'd have to keep it in mind.

We do have to keep in mind though that the information needed to pay the bill will always be of the highest quality. It makes sense, right? Like what are the pieces of our daily life that we take -- that we spend the most time on and these are things that really matter financially. Well, an insurance company is no different and as a former project officer of ResDAC to say, "Medicare is just a tiny little insurance company that makes its data available to researchers." So from an insurance company perspective, the quality of elements that they care most about at CMS are those elements that they need to pay attention to, too decide what is paid and how to manage the program, that's their primary responsibility.

So, Medicare data, administrative data in electronic format -- which is a huge help and they can help -- and they contain information about the services that are to be paid and should have been classic epidemiologic or bio stat, that would be sort of our numerator; and information about the people who are to receive the services which would be the denominator. And what's nice is that both are in electronic format, so they're easy to use. So what it means is we don't have to spend the time keying things

in and doing double entry and doing all of that stuff to get things from one format to another.

The other advantage of these data is that we can -- because they're electronic and because of the identifiers they contain, they can be combined with other sorts of information to make this dataset even richer. Barb will be talking about some of these opportunities tomorrow just before lunch, so we can talk about who's providing the care, who gets paid, which hospitals, which clinics, which physicians are doing it. The area of characteristics so that -- how many of you have used the ARF file, the area of resources file? Okay, so it's -- okay, so the area of resource file is a file that talks about -- tells us about geographic areas and their characteristics, what is their population, what is the density of providers. It is an easy link to do. And you can also -- other useful information whether it's cancer registries, or other characteristics about the provider -- of hospitals, you can create your own variables and these things can be put together to allow for a very nice, rich array of hypothesis to be tested.

So just to go through why use these? The first thing is, is really I think the most important which is clinical validity. So we're looking at healthcare and the answer is that these data contain information used by about -- services used by enrollees in the program. They have admission and discharge dates. These tend to be very accurate -- not 100% accurate because, of course, anything that you're doing with 43 million people you will find mistakes now and then, but the dates tend to be very good because this is an insurance company and the first check that they do when a bill is submitted is to figure out whether Medicare was responsible for this particular person on this particular day. So the dates are the first thing we check because if they're not under the Medicare program, then the bill gets rejected right off.

Diagnosis. A lot of procedures require specific diagnoses and so the diagnosis tend to be very good. Hospitals as we'll talk about tomorrow are paid based in part on diagnosis. The diagnosis tend to be very high quality.

Procedures. What was done is fundamentally the basis for how people are paid. They will tend to be accurate. There will be confusing parts about the coding, but in fact they will tend to be clinically correct.

And then the source of care; was it a hospital? Was it an emergency room? Was it a physician? Was it a nursing home? Was it a hospice? Because that is fundamentally where the check goes.

There's good demographic information. You'll notice on this slide, I say that the demographic information is largely reliable and valid. And I start with reliable first, and what that means as we all know, is you will get the same answer over and over again. So even if it's not perfect, at least it doesn't change.

One of the things that will happen when you see the claim forms that are

filled out by providers, there's a spot where it says name, date of birth, gender, race, all that stuff. And we see that on our own insurance forms. And so, what could happen which would be a huge problem would be if every time a particular provider got a piece of information right or wrong, it got transmitted, we would start to worry like, "Well, what is their birthday?" We've got seven different dates to worry about. CMS says the one on record from CMS which is from the Social Security Administration is the one that is used in all of these data. So we don't need to worry about errors happening from hospitals where date of birth isn't a major priority of theirs. We don't have to worry about those coming in and causing us major confusion.

So we've got really good information on demographics -- age and date of birth, gender, race, place of resident, and date of death. And date of death is again one that's very good and we'll talk about it later because it comes from the Social Security Administration. The Social Security Administration sends the social security checks. They don't like to pay dead people for retirement funds, so they tend to match these very quickly. I can also tell you from personal experience that there's now a really tight link between death certificates and all financial services. So when a death certificate gets issued with a social security number, the banks get notified and bank accounts will get frozen, social security will be notified, Medicare will be notified and it happens truly instantaneously. I was shocked at how fast everything got stopped. And, which is, I think ultimately about fraud but it -- what it means is that there are not going to be a lot of cases in the current era that are missed.

And then, just to remind everybody that we can link the enrollment and the use. So sometimes we have it where you can't put the two together and that limits us. So, for example, how many here have used the HCUP data, the healthcare utilization project data? So those are basically numerator data. They're hospital discharge data with no denominator to them. There's no at-risk population. You can approximate an at-risk population but you in fact don't exactly know who is eligible to be hospitalized in your state. The Medicare data, you know who is eligible, so you can do -- you have much more precision.

Population coverage. It's estimated that about 98% of adults, 65 and older are enrolled in Medicare. And over -- and Marshall and I a few years ago did a -- compared the national death index, number of deaths by age to the number of deaths we found in the 100% Medicare denominator file. And we found we were able to account for 99% of the deaths in the U.S. for people aged 65 and older and we actually did it, you know, we sort of stratified it and said we don't just want our total to lineup but are they lining up by age group, by gender? And we were frankly thrilled with how -- with sort of the level of coverage, so that the Medicare program really does end up covering everybody. So although we hear about reasons why people might be missed, in fact what we find is that everybody or just about everybody eventually figures out a way to be covered by the Medicare program. States can buy into the program, people can buy in, they can earn enough quarters, their children can buy in; so even people who don't quite seem to make the rules eventually seem to make it.

^M00:10:09

And then remember that and we'll talk about this, about 47 times this week that managed care -- so this is who is in Medicare program. There's a second level that we've got to worry about which is for what percentage of people will we actually find their utilization and that's where managed care is going to be a problem.

Medicare's program, we can study and I haven't -- I didn't update this for this time. I say 41 million; I think the number is now closer to 43 million people. The point is sort of -- doesn't matter. What it means is we've got a really massive sample size here. So it means even very, very rare conditions we will have enough people. We don't do studies in the Medicare program of 14 people. We do studies of 500,000 people and 4 million people. And if anything, what this -- the good news is that we can look at very detailed subgroups. So the first time I saw a study -- of somebody wanting to study healthcare and people age 90 and older, I thought they were ridiculous. I -- when I looked up how many people in the Medicare program were 90 and older, and I started to think about other things to look at for that subgroup, something we don't normally get.

So normally when we do our studies, we worry about power. We worry about having enough cases to find the difference that exists. If anything with the Medicare program, our problem is the opposite and that we will find that for every four days older, somebody who's 80 and older is, their chances of having something increases. That's crazy. I mean, nobody cares. It's too much precision for the type of thing we're looking at, and I personally have had arguments with journal editors where they've said, "I want p-values and confidence intervals" and I said, "Look, you know, I've got 40 million people here, you know." I've tried to argue, "We're a census, we don't need to do statistics." But I mean -- but seriously, the problem you will have often will be not low power, the problem will be more precision than make sense given what you're finding. So it's sort of public health significance rather than low power. So just be aware of that. But it means that you will always find a really nice p-value even on the things you don't care about.

Oh, you also have to be really careful with interaction terms for that reason. So we always have this rule of like if your interaction term is significant, you deal with it every interaction term is significant. So, again, you've got to watch it because the size provides opportunities, but it also provides -- it leads to challenges that you wouldn't normally have. These data are cost effective to use and I'm not an economist here so I just mean like cheap. So how many have ever like done a survey or a chart review study or any primary data collection? Okay. So if you've done it, you know how much time it takes even to just like ask somebody questions and get them written down and get them into a data file. Okay. Now, you imagine doing that for 41 million people, okay, and just imagine getting your demographic information. So do you think it would cost more than, say -- I don't even remember how much the file cost? And let's say more than \$10,000 to register 41 million people. Right. I mean, if the price -- primary data collection is incredibly expensive, incredibly time consuming and incredibly challenging.

And so this provides a whole lot of opportunities because we don't have those -- we don't -- the price per case is negligible, okay? We get access cross multiple providers, we get consistency in reporting format and that provides huge advantages to us. It means that we don't have to worry whether somebody was treated in a teaching hospital or in a rural hospital because they all chart -- they all use the same billing form. So we don't have to worry that if they were seen in an outpatient setting versus the inpatient setting that we're getting a different information. So what that means is, we can vary efficiently do our data analysis and get our data in shape.

I mentioned earlier about combining the data and here are some things we can do. You can combine the data with the census and the biggest challenge about combining Medicare data with the census -- to be completely honest is that the census keeps changing what data they release at which geographic unit. But, in fact, it's really -- once you can get the geographic units to match, it's really not hard at all. There have been numerous linkages of cancer registries with Medicare data. This year, Medicare is the most common, but also many individual states have linked their cancer registries with the Medicare data.

Other providers and insurance company, there's a program called VIREC. Is anybody here from the VA today? Okay. Sometimes we have VA people. So they take the VA administrative data and the Medicare administrative data and combine them, and the reason for that is that veterans who use the VA are not obligated to only use the VA. And so the VA was trying to understand the healthcare needs and use of their veterans knowing that it seemed really patchy and what we found -- what they've found and others have found, neither actually have been part of some of these, is that sometimes people go back and forth and they say, "Well, I do this at the VA and I do this at this hospital," sort of like shopping. "I get this at Costco, and I get this at Byerly's," well they do the same thing with healthcare across systems. And if you only measure one system, you don't really understand, you know, "How come these veterans, you know, they're such slow users of care. It doesn't make sense given the chronic illness they have. Is the VA not meeting their needs?" And the answer seems to be, "No." It's just that they use all of the resources available to them and if you only look at parts, you only see part.

There have been linkages with the national death index and with the state vital statistics. At various times, CMS has actually done the linkage for us and there are some periods of time where you can get cause of death. The challenge has been with that, that the states generally see cause of death information from death certificates as a revenue source. And if you can imagine paying \$5 per death certificate for the \$2 million -- you know, the two million deaths each year, it starts adding up. And so that's really been the problem. But it's been -- the policy has kind of shifted around so if you want to know cause of death, by all means, check with ResDAC to see for which years and for which people cause of death information is available.

The Minimum Data Set. This is information about nursing home residents

and it could be, again, combined with the Medicare administration administrative data very easily. There are surveys that have been done such as the health and retirement study which has been linked; then Medicare Current Beneficiary Survey is another that has been linked to claims; and then provider information and it can be both the Provider of Service file which you can get.

But I also have sometimes gone on and done stuff where I've looked up level one trauma centers, for example. I want to look at injury care instead of which injuries are treated in trauma centers and which ones are treated in community hospitals. And I just had one of my grad students look it up and manually code which hospitals were level one, level two and level three trauma centers, and you can do that. And the amount of time it takes is more about web searching than anything else. So, again, these are all very possible and very manageable.

So just to remind everybody that you can do these linkages at the group level based on geography, hospital and so on, but you can also take a cohort that you have and have it linked based on social security number or Medicare ID. So for example right now I'm working with the Women's Health Initiative, which some of you may know but is a large cohort study and randomized trial that's been going on since the early 90s and the Women's Health Initiative has been linked to Medicare. And so now what we have is we have all of the detailed clinical trial information, the detailed survey information, the biological information and we have their healthcare use. So we can look at things like what is the relationship between smoking status before illness and how well somebody recovers, or we actually know when something was measured relative to an event.

The data are available in a timely basis and this is actually I'd almost argue coming to be as much of a negative as a positive. So we always worry that our data are getting out of data, right, but now it's getting to a point where somebody says, "Well, you're only using 2009 data, what's going on?" And I'm thinking, "Like, do you have any idea how hard this is to get this all happening? And it's only a year and a half old. It's not that far out of date. Where normally if we did a survey or something else there would be a lot more tolerance for data being a few years old. It's fantastic if we're looking at the impact of policies.

So there are policy changes and it used to be that we'd have to wait five, 10 years to figure out what happened and whether a policy was effective. And we can now start monitoring those policies in real time. So that's fantastic, but for a lot of our other work, it's kind of a pain. But, again, we don't want to complain too much.

^M00:19:47

So --okay, so like what's the downside, right? The biggest thing is that this is a record of care received. Okay? So if nobody diagnoses you with somebody, you don't have it. Okay? Hypertension, depression, diabetes, osteoporosis are often underdiagnosed in clinical settings and if they were never diagnosed, they didn't happen. Nobody goes back and fixes the record, so nobody goes and says, "Oops. We missed that one three years

ago." They really have been hypertensive the whole time. We find out when it was clinically recognized, not when it started. So that gets to be challenging.

If you think about, how many of you have watched the show, "House?" Okay? Now, imagine House's claims data. Okay? Is it hypertension? No, it's a brain tumor. No, it's an infection that he got in the Amazon. No, it's a broken leg. Right? And every one of those things would show up. Now, fortunately, he's just a TV show but you will see stuff like that where you're kind of looking at this string of diagnoses trying to figure out, "So do they have osteoporosis? Should I count them as a "Yes" or a "No"?" And a lot of our algorithm time is really spent trying to figure that stuff out. Unfortunately, they're a minority, but those are the challenges because, again, they never go back and say, "Oh, they figured it out, broken leg," and then erased all the old stuff. It will all be there in the record and we need to figure out a way to work through it.

It's a record of care received, not care needed. For example, with cancer one of our challenges is finding recurrence. We can find treated recurrences; we cannot find untreated recurrences. Okay? Clinically, there are still recurrences, but if somebody chooses not to get treated, we don't see it. If somebody is -- has a consult and then doesn't get the treatment, we don't know whether it was ever suggested or not. So in my work I look at use of hospice prior to death but I don't know who actually was given the option of using hospice versus who wasn't. I just know who ended up there at the end of this whole process that I can't see.

We have this trouble when we look at some of our disparity's work. So we look at differences in treatment patterns across age groups, across racial groups, across ethnic groups, across socioeconomic groups and we can see that the treatment is different, but we don't know why. Was it because nobody suggested it? Was it because they turned it down? Was it because they couldn't get an appointment that they wanted it? Like it was suggested they wanted it and they couldn't get in? So we don't -- what we know is what happened. We don't know what somebody tried that happened, what somebody suggested, what could've happened?

And then there is some things around denied services. So what it means is that if you're looking at a new technology, you need just take the time to look up to make sure that Medicare actually covers it. So there are services that Medicare doesn't cover and there are services that Medicare covers with restrictions. The coverage policies will change over time, so there'll be maybe a new technology where CMS says, "You know, there's not sufficient evidence that this -- that this treatment works. We will not pay for it." And then as the scientific evidence accumulates, they will change their policy.

If you want to look at use of that service prior to coverage decision, you won't find it. We will get denied claims, and so one of the things that sometimes people do is they say, "Well, why don't I just use denied claims to find it?" And the problem there is is that if a provider knows that a particular technology is not covered, in some cases they will submit the

bill anyway and in some cases they won't, and that makes it a real problem.

So years ago in the '90s before screening mammography was a covered benefit, people were trying to look at mammograms in Medicare. The -- a diagnostic mammogram has always been covered by Medicare, screening mammograms were not. But the first paper looking at screening mammography came out before there was a screening mammography benefit. So people will analyze things that in fact are sort of wrong knowing the program. But I asked a friend of mine, I said, "So do you submit -- does your office submit bills for screening mammograms to Medicare?" And her answer was "Yes." And I said, "Well, why do you do that? You don't get paid." And she said, "Because they ought to have that benefit and so I submit the bills so that every time they see one of these they get reminded that they ought to have the benefit." And so this was somebody who very explicitly said, "I know the rule. I know it's going to be kicked back. I'm going to do it anyway." Other people have said, "Well, I know they have secondary insurance and the only way the secondary insurance will kick in is if the Medicare rejects it, so we need to do it. And other people have said, "We don't do it. It's just a waste of time because we know it's going to get kicked back. There's no benefit to us to doing it so unless somebody asks, we don't."

So there's a huge range of how denied services are handled by hospitals and clinics and providers. So it means that if you're looking at new technology, if you're looking at services where there's a controversy, you need to take some time to do your legwork around the policy and make sure you understand what it is you're going to be seeing.

Diagnosis codes don't always tell us everything we want to know. Cancer diagnosis code -- cancers for example, are coded by anatomy, not by anything that I would argue we really care about. So lung cancer is coded as a 162.xx and the X is referred to left lobe, right lobe, upper or middle, lower; but they don't tell us whether it was metastatic or local or small cell or non-small cell or high grade or recurrent, or any of those things that, frankly, we'd want to know to have a sort of a clinically meaningful cohort. We don't get that. Hypertension is coded as hypertension, but we don't -- I mean, what we really want to know is what is their blood pressure, how long is -- how long have they had it? Is it controlled? What medication are they on? Right? And so we've got to watch it because we don't always see enough to really get a tight risk cohort.

Congestive heart failure is actually one that's improving a little bit. It used to be there was a single congestive heart failure code, 428, and now they've expanded it so that the XXs divide people up into sort of degrees of heart failure. But it used to be that you could see somebody who had heart failure for five years and you didn't know whether it was stable, or getting worse.

Diabetes is another one. We know if somebody is diabetic but we don't know how well controlled it is. We know if somebody has a hemoglobin A1C test, but we don't know why that test was ordered. So these are -- so sometimes we have -- we have enough information to do some stuff but not enough to do more.

Many drugs and procedures have multiple indications and that's the problems. We may know that somebody received a certain treatment but we don't know why. And just with that reminder, there are no diagnoses on the Part D data, which is the pharmacy data. So we know what somebody gets, we don't know why they got it. So if -- so again, it's piecing together the bits of information from what their clinic visits that they had and what the pharmacy said somebody was given and trying to put that together into a coherent story.

Just to sort of say really strong, I would really strongly discourage blanket statements like, "The only reason anybody would ever do X procedure is for X condition." In my experience, there are very, very, very few absolutes like that. And so just remember, we cannot tell intent, cannot tell why something was done. We cannot -- do not assume that decision making was always based on best practices. Well the only reason anybody would ever do this is because, you know, of this assumption, and in fact, there's an awful lot of crazy care that happens.

Different care settings use different coding systems for procedures. So we've got diagnosis codes which are easy. Everybody uses ICD-9 diagnosis codes. I -- sometimes I get asked about switch to ICD-10 which is the new coding system. Death certificates have already switched. There are discussions about moving healthcare. There are dates. The dates have passed. I personally think when it was going to happen we will hear about it because it will be a major overhaul that affects every level of the healthcare system, and I think it's going to be longer than we think because it's going to end up being a very expensive shift. So in the short run at least, I wouldn't worry about a shift to ICD-10.

Procedure codes are a problem. Inpatient care is used -- is coded using ICD-9 procedure codes; and these procedure codes are 4-digit codes and they refer to big things that are done in hospitals like open heart surgery, radical mastectomy, hemicolectomy -- big jobs; hip replacement, partial hip replacement, total hip replacement. And that's about the level of detail you get because that's about how the level of detail that matters to a hospital.

^M00:30:12

The carrier code in the Durable Medical Equipment Code are using CPT codes which are created for -- the CPTs are created by the AMA and then what are sometimes called "HCPCS or level 2 HCPCS" are created by CMS to fill in the things that are not included in the CPT codes, and those are five digit and they are much more detailed.

And to give an example, with hip replacement -- with partial hip replacement, there's one ICD-9 code. There actually are two procedures that can be done. There's two different devices that can be used to treat that hip fracture -- nails and screws, if I remember right. They have different CPT codes. So the technical work that the surgeon does is different depending on which device they use. Hospital doesn't care. In the hospital data, we can't tell which device was used because it doesn't matter. We have to look at the surgeons if we want to know what exactly

was done. And so often when I am able to talk to clinical colleagues, we will look it up and we will say, "You know, for the surgeon it matters very much whether it's laparoscopic or whatever. There'll be details that will matter that won't be shown in the ICD-9 procedure codes.

So -- but inpatient care, four digit, big picture stuff, carrier and DME, CPT codes, 5 digits, far more detailed and hospital outpatient care. So this is care provided in a hospital outpatient setting, so non-admitted; it could be emergency room, it could be a clinic, it could be radiology, it could be day surgery; it's coded as a mix of CPT and revenue center codes. Okay? So hospitals have to do both. If you are looking at very, very old Medicare data, if you have data going back to the '90s, it gets trickier because hospital outpatient used to use ICD-9 procedure codes and then they migrated to CPT codes. So you just have to be aware that this is where things get hard as you're looking at whether the same care was provided across settings.

And then there is not a terrific crosswalk between ICD procedure codes and CPT codes. So it isn't just that these two CPT codes map to this one ICD-9 procedure code. Sometimes it's -- well, this one maps to this, but this code could either be here or here depending on how it was done. So it can get very tricky and in some cases we're forced to make less than optimal decisions because it's all we can do, but just be aware of it, and think about it and spend your time pouring over the books and you'll sort it out.

There is limited clinical information, so we don't have physiology, we don't know blood pressure, we don't know pulse, we don't know ejection fraction, we don't know body mass index. We don't know some really useful things. Test results are not included, so we know whether somebody had a PSA test or an angiography, whether they had a surgical specimen taken; but we don't know the results.

A few years ago, somebody published a paper that said, "If you see a PSA test and you don't see anything after that, you should assume the test was negative." My personal read on it is if you see a PSA test, just for prostate cancer and you don't see anything after it, you should assume that nothing happened after that test. It may be because the test was normal and it may be because they -- the test was abnormal but they decided that it wasn't abnormal enough to follow up, or it may be because they never notified the patient, or it may be because of any number of things. And that's sort of the point. You cannot tell from the existence of a test what the test results were. And that gets very frustrating as we're looking at quality, so if anyone here studies diabetes, one of our quality metrics is studying lipid levels, blood pressure control and hemoglobin A1C control. Right? And we can find the lipid test in the claims and we can find the hemoglobin A1C test in the claims, but we cannot find the results. So we know if somebody's being monitored, but we don't know whether they're doing well. I mean, the tests aren't magic. The tests are just a way to meet some end and we can't tell usually whether that end was met.

Yes?

>> Have you started using the PQRS data yet?

>> Beth Virnig: The question was, "Have we started using the PQRS data?" The PQRS data are data that are being collected on provider's quality in order to actually assess this. And the answer is I have not -- some of my colleagues do, and one of the things that I hope will come out of these data. So these are data that will have services and test results in one place. And so one of the questions is we may be able to move forward, do some algorithm building, but again, it's going to be these small studies that are actually going to let us do this stuff. It's not that small but compared to the whole Medicare population.

And then exact timing isn't noted, and in some cases -- so for example, if there are things that are really time sensitive, treatment of heart attacks and strokes in the emergency room, timing is very, very important for clinical quality. We cannot do this. We cannot tell it. Emergency room visits, you can tell if somebody was admitted on a Saturday but you can't tell if they were admitted at 4:00 in the morning, or noon, or in the middle of the night. So, again, in some cases the timing would be very useful, we cannot tell that. So if you're looking at questions like, "Well, are people going to the emergency room because their physicians' offices are closed?" We can't tell which Monday after -- which Monday emergency room visits were during typical office hours and which were not, so some of those studies are just going to have to be done using other data sources.

The data sort of in summary are limited to covered benefits for which claims are submitted. Prior to Part D, Medicare had no pharmacy benefit, so outpatient medications could not be studied at all. With Part D, studies will have to take formularies into account. Formularies have not consistently been available throughout the history of Part D, so the question as always will come up before the formularies which are 2010 of, "Is it that this person did not get drug X because they didn't want it, or because it wasn't one of the choices in the plan they picked?" So once you get a formulary, then at least you know whether it was in the list of options. But that's always been one of the questions when there is sort of a range of choices is, is it that they chose the step 1 drug and stayed with it because it was good, or because there was no step 2 drug available for them?

Covered services for which claims are not submitted are not included. So there are flu shots that are a classic problem here. So grocery stores will have flu shot clinics. You can go into a pharmacy and get your flu shot, our health department runs it, and if this -- if the claims aren't submitted, we won't know about it. And I actually -- this was sort of brought home for me a couple of years ago when my mother was talking about her mother and she said, "Well, I went up to get her her flu shot and she didn't have her card with her. And so we thought about driving back to her house which was 10 miles and driving back, and so I said, "How much is it?" And they said, "\$6." And I said, "Here." And we just paid cash for it because the price of driving home, getting the Medicare card and driving back was higher than six bucks, in which case -- so if you could imagine, if we were looking at her and trying to assess whether this 96 year old was up to date on all of her preventive care, we would've classified her as not being up to date. Right? But in fact she was, just that no

bill was ever submitted for it. So we just have to remember this and that's why when -- people who are studying flu shots, Marshall is -- this is one of his areas of research, so I encourage you to ask him about it tomorrow. But when you look at the claims, the coverage from Medicare is almost always lower than other estimates of flu shot coverage for the elderly, and this is the likely explanation. And then as we said earlier, some services just aren't covered, and if they're not covered they won't be there.

Managed care. There's going to be no information about Part D services for managed care enrollees, and there's little information about hospitalizations for managed care enrollees. The hospitalization data has been mandated to be reported since about 1990. It has never been released for researchers. Some friends of mine who are researchers at CMS have reported that when they have tried to do validation studies, they have been unhappy with the results that they have seen. So I think although nobody quite understands why there's a fair amount of concern that there's some incompleteness or some inaccuracies in what they are getting and that's the basis largely for the reluctance to release it.

^M00:39:58

And then just to remind everybody that not all beneficiaries have Part D coverage drug coverage and not all people with Part D or equivalent coverage will actually have claims in the formulary in the files -- in the event file. So the event files will be about 50% of your population. Okay. So if you've got a population of 100,000 people, I would expect based on my experience that about 50,000 of them will be in the Part D event file. So you can see, I have not done the work. It's frankly on my list to actually ask the question of whether the people who are in the event file are systematically different than the people who are not. We have reasons to expect that they are, but in fact that's going to be one of those questions when we start talking about bias and confounding. Are we seeing a biased subsample of people using the Part D data?

Variable quality. As we said -- as I said earlier, payment is a big deal. My rule is, if a field impacts payment, the quality will be better than if the field doesn't. What does that mean? It means that different types of care have different payment rules and it means that even if you're not an economist, even if you don't particularly like dealing with money, you're going to have to spend a little bit of time looking at the Medicare coverage rules and understanding the payment rules and how things are handled.

A lot of the confusion that I personally have found and that I have found with researchers over the years, happens because we don't understand how CMS manages the Medicare program; and that when we take the time to actually figure it out, it actually makes sense. I mean, some of the rules you're like, "Well, I don't understand how they got that rule," but once you understand the rule, in general, the patterns we see will make a lot more sense.

And there are even crazy examples like when policies from CMS are based on the federal fiscal year and not on the calendar year. So somebody called me up and said, "Well, they said this changed in 2009, how come I'm seeing

it in 2008? Right? Now, that doesn't make any sense." Well, the answer is, it was changed the 2009 federal fiscal year and you saw it on '08 because you saw it in October, November, December, which is part of the fiscal year. Okay? So those are the sorts of things that if you take the time to understand the program, it's going to make a lot more sense.

Other implications of the rule though, comorbidity and severity of illness information may be inconsistently reported -- recorded in the claims, because not all settings of care -- it doesn't matter how severely somebody is ill. So knowing that, we'll tell you where you expect it.

And in some -- not all components of treatments may be included in bills. And this is -- this is one we've had debates about. If the payment is very, very low for something, at what point is it no longer worth building for? And if you hit that point, well you say, "Well, this is only worth 42 cents and it's going to cost a dollar to send the bill out." Some providers may choose not to bill for that 42 cents. And so you sometimes will say, "Well, these three things ought to show up and I'm only seeing two of them." My personal experience would be the one you're missing is going to be the one that costs 42 cents and that they rarely miss the \$10,000 procedure.

When we think about data quality, the data elements provided by CMS will provide consistent information. You'll get this answer over and over and over again. The record layouts are very important. They're not just for your analysts. They're not just for the person reading in the data. They will contain important information about the assumptions, data combinations, rules, where the data came from and that you should keep them handy. We keep copies -- every person on our study will have their own copy of them. We now, for the most part, they're electronic. I still like my big thick binders. They're just -- it's a generational thing. But just be aware that they're very useful and you should have them -- everybody at every level of your study should have them -- have access to them.

So here are some things, and some of these are pretty old, but I think they're still timely of things that could only be done with claims. We've all seen these maps. These ones are from the Dartmouth Atlas where people have done U.S. maps variation.

This is coronary artery bypass procedures. You'll notice it's like almost 20 years old, but the point of it is, these maps happened because of administrative data. We cannot get this level of coverage using primary data collection, it just doesn't happen.

And you can see this one. I still like this map a lot. This is knee replacement surgery, and the reason I like it -- and I'm going to get in trouble because I'm going to walk now. This is the problem with videotaping, is you see this darks is that there's a ton of knee replacements going on in the upper Midwest, right? It's probably though the snow and ice, and what I find interesting is that you also see it down here in Arizona and you see it on the west coast of Florida, and you see it right here in this area of Texas, which is where all the Minnesotans go in the winter when they get tired of the snow and ice. So that is my theory of these

patterns. Other people will look at these patterns and say, "No, it probably has more to do with the density of orthopedic surgeons," or something like that and not blame it on the ice. But the point is, is we can ask those questions and we can dig in and we can start looking at these maps in a way we never could before.

We've all seen the volume outcome relationship studies, do high volume centers do better than low volume centers, do high volume surgeons? Is it practice makes perfect? Or is it sort of a selection thing where the people who are the healthiest can travel the farthest. Right? If you're healthy enough to travel to the Cleveland Clinic, you're probably a pretty good operative risk versus the person who gets taken by ambulance five miles and everyone's nervous the whole time. Those studies, we've been able to do using administrative data in a way we could never do before.

The studies are variation in treatment patterns. Why is it that people -- some people are treated this way and some people are treated this way? And what happens because of it? Those studies have all been facilitated by Medicare data. And then we can ask even sort of much more socially relevant questions like, "Do racial differences and outcomes persist after treatment patterns?" Is it that there's something different about racial groups? Or is it that once you take into -- once you treat everybody the same way, they in fact act -- they in fact have the same experience.

In the late '80s, there was an editorial in JAMA or the New England Journal called the "Yentl syndrome," and I don't know how many of you remember that old Barbra Streisand movie, "Yentl." Anyone?

Anyway it was about this woman in 1800s Germany who wanted to study but girls weren't allowed to go to school. So she cut her hair, put on boys' clothes and was allowed into the school and did as well as any boy. And the whole point of the article was about cardiovascular disease saying -- arguing at the time that a lot of the variation in survival after heart attacks was not because the anatomy of women was different and they had small vessels which were sort of what the surgeons were saying. But it was in fact because women weren't being treated like heart attack patients; and that if you treated them like cardiac patients, they in fact responded like cardiac patients and they did quite well. That debate probably is still brewing some ways and versus others but the point of it is, with data files like these, we can actually understand it and we can look at the minutia of how somebody was treated, where somebody was treated, the timing between events and try to understand that.

So getting ready to use the data. So my thing with these data, these data are hard to use and I say that now. I've used -- been using them for almost 20 years and they're still hard. And the reason they're hard is because they're so easy technically. You know, you put them in, you do a -- you say, "Okay. Give me a crosstab on the following variables." And what you really need to be doing to be successful using these data is sort of stop and think about it and plan it out, and make sure you've really gone through everything before you start running data, and that's my big advice.

So when I -- so the good news about the process of writing grants to fund this is we're forced to some extent to do some of that. Those are 12-page grant proposal probably allows us to back off a little bit more than the 25-page did because you have to do something to fill up those 25 pages, so we tended to write out a lot more detail.

But sort of the questions would be like, "Who is in my study?" Like really basic question but like what are my criteria? Are these really the right criteria? Is there anything else I want to include? Anyone else I want to get rid of?

What are the events? What is it that I am measuring and how will I measure it and how will I be sure that what I am measuring in fact is what I want to measure?

And then finally, what are the key adjusting variables? What are my covariates? What are my confounders and how will I find them? And how confident am I that what I think I am measuring is what I am measuring?

^M00:49:56

Taking the time to think through those things will save you months of time in the end, because eventually -- in my experience at least, you will finally get to this point, and the question is just whether you do it first or you do it at the end. Eventually, you will do it.

So, you know, sort of think about your target population, and then think about things like if you're pulling from a denominator, everybody in your denominator should be able to have events. So what that means is it means that if you're looking at medications, you want to restrict your denominator to those people who have the type of Part D coverage that will show up in the Part D events file, because if they don't, you're going to underestimate medication use because they're not eligible to be in your numerator.

Sometimes people will do things where they won't put their numerators and denominators together. They'll keep the two files separate, which is actually easier at one level, right, because we've got two smaller files. And the problem then is that sometimes you'll end up with people in the numerator who aren't in the denominator. So it can go either way and it's really about thinking carefully about who is it we're studying and how do I know that those people are actually eligible?

Now I know there is sort of a group of epidemiologists who will occasionally argue about, "Can you have men as controls in ovarian cancer studies?" And we're not going to get into that. My advice would just be make sure that everybody in your denominator can actually have the events you want. And there's some really basic questions, so like we talked about geography and so they said, "Well, let's look at location of residence." And this sounds really easy, right? Like I want to look at a cohort -- in fact I did this years ago; it's down in Florida and like, "Well, let's look at a five-year cohort of people residing in South Florida." Cool, right? How hard is it? I mean, that's not that hard right? Well, when we did our study, we didn't really think about it as much as we in retrospect should've,

and we pulled everybody who had lived in Florida for that five-year period. And what we ended up with was a population of really healthy old people. Right? Because by requiring continuous five-year residence, it means that anyone who died, we've got rid of. Right? So now we've got a healthier group, and anybody who would have aged into the program, they weren't eligible, so now they're older. So now they're healthy and they're old and so -- and so then it made us realize that sometimes we've got to look at our study design and that we don't always want to do a cohort study that sort of always seems the most obvious.

But in this case, for our study, we realize we really needed to do a panel study. We needed to do five panels in order to really characterize the population of South Florida over a five-year period. We needed to change a little bit how we were conceptualizing our design. We figured it out because we started -- you know, because we're like, "Well, how can we have a mean age of 75?" You know, and just like it's too high. So, again, these seem like really simple things to think about and they certainly were once we figured out our mistake, we exactly understood what we did wrong. But the goal here is that maybe we can avoid having you guys make that mistake, too.

Diagnosis are a real challenge, diagnosis and procedures, because what will often happen is somebody will say, you know, I do not have stock in Google. So google "ICD-9 diabetes" and some -- and it'll get back 250. "Cool, this is it. See, it's not that hard." And then what happens is if you actually look at the code book or you spend a little bit more, you realize there about 14 different codes that in fact can signal that somebody's got diabetes. So what happens is, we'll sometimes have it where we'll think we have all of the codes and then we'll realize we've missed one or we've missed a couple. And that can cause real trouble. It causes trouble in a couple of ways. The hardest way would be if you asked for your data from CMS and you said, "I want only cases with this code 250," and then you figure out after your data come, then it's financial trouble.

But for those of you who work with an analyst, I can tell you that when you're working with really big files and you say, "I want 250s." And they say, "Are you sure?" And you say, "Yes, I'm sure." And they're like, "Are you really sure?" And you're like, "I am really sure." Go back and forth you finally say, "I know what I'm doing." And then you've got to walk down the hall two days later and say, "Remember that work you just did? Well, guess what? I forgot the code, so you need to redo it and please don't kill me." But seriously, I've done that with my analyst like is -- she usually checks my work now because she's had to redo it so many times. But it can take a huge amount of time to have to go back to the raw data and add another code -- and yes.

So I think the suggestion was, the variety of ways you can do it, one is to look at a medical coding guide and that sometimes they will have contingencies. I would also say invest in an ICD-9 code book, okay? You can either get an electronic one or you can get a paper one. We got ours off eBay for like \$4 each. We used coding. I mean, your coding book can be a year old, right? I mean, our data are, you know. Right? So the

hospitals will get rid of all of their code books and then you just pick them up. But take the time to look at the code book, like open the book, find the section, read what it says because it will also often point you to other places, you know. This definition does not include this which is coded here. And look at the index and look at it a couple of different ways. So there certainly are strategies, but the strategy has to be -- you've got to spend -- be willing to spend a little bit of time because you can get into a lot of trouble.

Yes?

So it's a good point. So CPT codes in particular will change; and they change -- and sometimes dramatically. There was in about 2007, all of the breast cancer surgical codes changed. And so I have a colleague who'd done all of -- she had done everything correctly, right? She'd done all of her legwork, she'd looked at all of the code books, she requested her data for a five-year period and halfway through the fifth year, there were no cases. And what happened is there were changes in codes that she didn't catch at the time she put in her data request, so she ended up with a four-year study and it ended up being sort of okay. But the point of it is, codes change, the changes are more likely to happen in procedure codes. So there's two pieces. With diagnosis codes, things can often show up in a variety of places. Procedure codes are easier to find. They tend to be more specific but they also tend to change more, so you've got to really take your time to do the legwork.

If you look at current policies, which is one of those easy things to do, you've got to remember that those are current policies for current coders. So, when I look up, what the current coders are using for a procedure, there's no guarantee that my data which may be from three years earlier will be using that same set of codes. And in general, like when you look at insurance companies, Aetna's really good about having coverage decisions and codes and they will have everything really current. But we have to remember that we usually want things that are a couple of years out of date and those can be harder to find, so it means you're going to have to be prepared to do some checking and some logic checks. Like if I'm right, this should be happening. Am I saying the thing -- am I seeing things that actually make sense and that's it. So there should be an underlying assumption that the data actually makes sense.

When you're looking at your population, you know, think about, do you want to define it or limit based on demographics, on specific coverage? Do we want to limit our study to people with Part A and Part B, no managed care? Do we want to limit it to state buy-in? What do we want to do? Do we want to limit -- do we want to get rid of the disabled or the end-stage renal disease, or do we want to include them? These are all key decisions that should be made ideally before you start your studying.

And then -- then there's some easy ones like membership in a defined cohort. When we think about the outcomes of interest, the easy one is a receipt of a diagnosis or a particular procedure. Questions happen like when is somebody diagnosed with something? Is it the first time it shows up or is it when the confirmatory diagnosis shows up? So if you see somebody

for a chronic disease and there's one code of diabetes, is that the day they became diabetic? Or do you want to wait and make sure that they actually are called diabetic a second time? With healthcare use, just remember, there's a lot of options.

Hospitalizations are easy but readmissions are also incredibly useful. Look at clinic visits. Look at duration of care or time until an event. And time is one of these opportunities that I'd encourage you to use. It's a different dimension and in some cases what I found is that at the end of five years, the event rates may be relatively similar between two groups. But in fact, how they get to the five years is maybe quite different. So doing a Kaplan-Meier-type approach, Koch's models, time to event can sometimes provide different insights about the trajectory and experience of patients.

Money is always a useful measure. It's a good summary but also in some cases it really helps us understand patterns. You know, what are the incentives? What are the economic incentives underlying these patterns and there's a series of studies and there's always room for more asking whether clinical -- whether clinical factors or financial incentives better predict use of specific services. And again, with these data you will have that information.

^M01:00:02

When you're studying time trends, there's some -- it's -- you know, often when we do these time studies, we do it because there's a technology or there was an important paper, a report that came out and we want to see the effect of that. So what happened after this report came out saying that this was -- this technology didn't work. And that's sort of -- those are the fun sort of health services research studies. The problem is, is that in some cases, other stuff is going on that we need to keep our eye on to make sure that we really are getting the right cause-and-effect relationship.

So there have been major payment policy changes. I'm going to point out one of the most obvious ones, which was DRGs in 1983 which completely overhauled how hospitals were paid. DRGs have been adjusted since then. So with DRGs, it wasn't just about payment but it was about what factors go into payment and where the incentives were about how hospitals code and describe the people that they are taking care of.

Medicare-managed care. Marshall talked about that this morning; I'll be talking about it again this afternoon. The managed care program started. It has been continuous, but if you look at enrollment, it looks like a rollercoaster; it's gone up, it's gone down, it's flattened. Every time we do that, we lose people from our population or gain people. And so we've got to watch to make sure that our time trends in fact make sense and that we're not just seeing something else. There have been changes in the coverage of preventive services, the expansion of Part D. There have been changes in payment policies, changes in CPT codes and all of those things make it challenging to look at something over time.

If you are watching a time trend, it also looks like something just fell off a cliff, you'll see this would be going along and all of a sudden it's just gone. That's usually a sign that there's been a coding change, and one of the things you want to do sometimes is we will sometimes just do a crosstab of year by the various codes that are in our algorithm to see if there's a code that disappears. And once you know their code disappeared, then you know -- you've got to go look for another one and figure out what happened; but, again, you will certainly see this. So you'll see codes getting substituted, CMS having a ruling or an advice saying, "We don't really think this is the right code. We think this is the right one." And even though the code may still exist in the code book, people will stop using it. Time trends are tricky because it's rarely just one thing that is moving.

When you look at adjusting variables, we can look at demographics, we can look at location of residence and comorbidities. And fundamentally, when you're doing the selecting for your study, what you want to do when you put in your request is you want to sample wide. You want to take the time so you can say, "I'm going to ask for the biggest cohort that I can justify that has every possible way the thing that I want will show up." Okay? And then what you need to do though, because if you forget something you're going to have to go back and make a second request and you'll have to repay, you have to go through all of this paperwork. And even though you will get permission from CMS pretty quickly, you say, "I forgot a code, can I get permission to add one more code?" They're going to say, "Sure." That's not the issue. The issue is paying for it.

So the ideas that you're better off sampling wide but then you have to analyze narrow. So then what happens sometimes is people say, "Well, I've got all these data. I don't really want to exclude half of the people I've got, so I'm just going to keep everybody." But then you end up with something that's not particularly well defined either, because you end up with all of the stuff that isn't really the direct focus of your study. So it's about hitting that balance and being willing to sort of think about, "What is every single possible way I could find this?" And then taking the time once you get the data to say, "Now, I'm going to figure out who exactly I really care about," and focusing your analysis on that group. And by doing that -- I mean, otherwise you just end up with this mess, right? So if you don't do the careful analysis and careful focus, your study won't make any sense. And, again, with data files this big, you really -- power should not be a concern.

And with this, with all of these steps, I will tell you sort of my group's personal thing, I have a three-paper rule, and that is to say that whenever we do a study -- whenever we do a data poll, right, which means we take the data, we compile it into analytic files, we try to make sure that we have three papers in mind that we could do from that population we have pulled. Because it takes a lot of time to get this done and to get this into an analyzable format that if we don't have multiple things we can do with it, the price to do that one paper is really pretty high. It's way too high in my opinion. So, and I'm not talking like sausage slicing here, I'm talking like three papers like what is the incidence of some disease,

what are the variables predicting cost of treatment in the first year? What are -- what is the relationship between race and comorbidity and intensity of treatment? I mean, those are three distinct legitimate papers you could do on a topic. They will all use the same basic analytic file and that's why we do it that way because to create the analytic file is a lot of work even for a very, very experienced analyst. So have a plan and have a big picture in mind when you do this, and that you will find yourselves a whole lot more successful and a whole lot more productive because, again, this upfront stuff takes an awful lot of time. And that is it.