

## Linking Inpatient Registries

*Dr. Leslie Curtis, Ph.D.*

*Dr. Zubin Eapen, MD*

Thank you very much. It's a pleasure to be here today.

Today I'll talk about linking inpatient registries with Part D data to assess discharge adherence. And I have no disclosures to report.

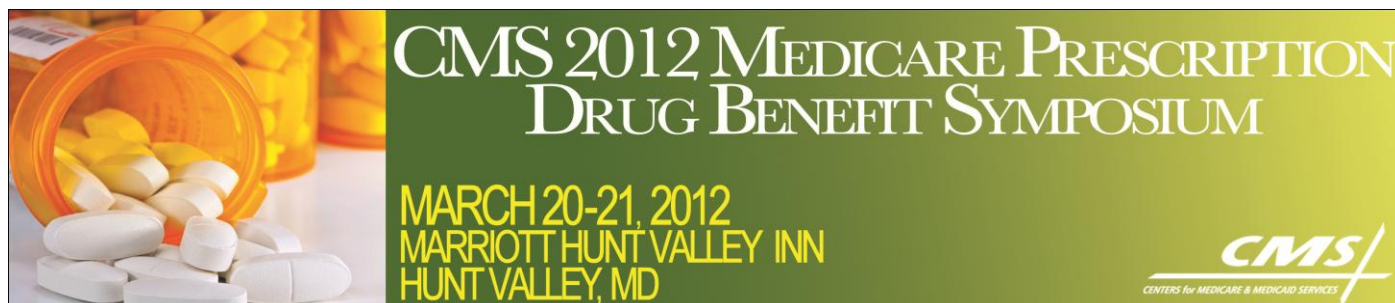
The learning objectives for this talk are to – to first demonstrate how inpatient clinical registries can be linked with Medicare claims data and eventually with Medicare Part D prescription drug event data. And then to illustrate how the linked data set, that is Part D linked with national registries, can be used to assess post-discharge medication adherence.

Over the last couple of decades there really have been many, many data sets that have emerged with the potential information that can be linked to – to Medicare claims data. This slide shows just a few both clinical registries and epidemiological cohort-type studies. Whenever the question comes about, you know, can we link these data with claims data in order to assess post-discharge kinds of follow-up events, the next question is always, well how might we go ahead and do that. the most common approach, and one that's been used with quite succ- quite some success with Seer Medicare and other examples is to use direct identifiers, like social security number, date of birth, sex, address, and the goal there is to identify each patient, or each participant in the Medicare denominator file essentially.

We have some experience at Duke linking Medicare data with longitudinal epicohort studies from the National Heart, Lung and Blood Institute. I've listed a few of those here. And this approach, in doing this we've used that direct identifier approach.

Now the inpatient registries that we also house at the DCRI are different in that they don't typically include social security numbers and those direct identifiers that we would need to identify a specific participant or individual in the Medicare denominator file. What we have instead are what we term a series of indirect identifiers. Service dates. We often have date of birth or age. And we have the sex of the individual. So the goal shifts a little bit and becomes to identify each registry hospitalization in the Medicare data. We have many examples of where we've done this. I've shown a few on this slide. To date we've linked the American Heart Association's Get With the Guidelines registries with Medicare claims data, the National Cardiovascular Data registries with claims data, as well as some other heart failure registries with claims data.

So how exactly do we do this? Well, consider a hospitalized Medicare beneficiary. And on this on this slide the stripe – the vertical – or the horizontal bar shows you some sort of duration, and each of the vertical bars denotes a hospitalization, let's say. Imagine that that Medicare beneficiary is – is hospitalized a second time, and that hospitalization takes place at a site where they also participate in perhaps a national heart failure registry. So at that point data are collected on that – on that individual for the registry, rich clinical detail about symptoms, signs at presentation, the course of the hospitalization, detailed clinical data essentially. Now if we have a way to link that registry hospitalization with the



Medicare claims data, we could then also have the potential to gather additional end points from the claims data. Some of those are listed below. They would include mortality, readmission, certainly subsequent procedures and potentially adverse events.

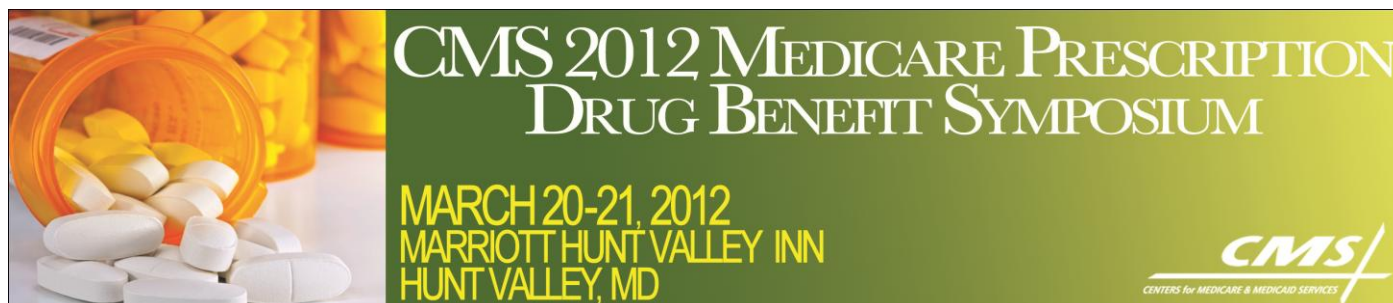
So how might we do this using indirect identifiers? Now let me take a minute to orient you to this slide because I'll show you a series of slides that look similar but – but they're important in some – or they're different in some important ways. Here across the – the X axis, we're showing the percent of unique records. And by that I mean what proportion of records are uniquely identified by the variable that we list in the upper left-hand corner. In this date, admission date. And the scatter plot then along the Y axis shows the number of records within a given site. The number of heart failure records within a given site. We can think of the site as hospital here. So, not surprising, you look in the 100% sample from 2007 of Medicare heart failure discharge records, we see that admission date is not a particularly unique identifier. Especially in large – in large volume centers, right? And that stands to reason. But what happens if we add additional information? What if we use, for example, admission and discharge date? Well, with that combination we see that our line begins to shift toward the right telling us that when we have admission date and discharge date together with the site, we begin to uniquely identify that hospitalization for a given individual.

What happens if we add date of birth? And the answer is we achieve an almost unique identifier irrespective of the size of the site at that individual level.

Now I'll show you what happens if we don't have discharge date, it turns out that the date of birth is really, really a key bit of identifying information here. Importantly not all – not all registries collect all three elements of the date of birth. Some will just collect month and day, for example, and in that case we do fairly well with that as well. Other registries don't include date of birth at all but they do include age at the time of admission, and again, we do fairly well in terms of identifying unique hospitalizations on the basis of these three bits of information. When we add sex to the mix, again, our ability to distinguish unique hospitalizations improves. And then here we just show what happens if we allow a little bit of fudge on the admission date or a little bit of fudge on the age of the beneficiary.

Now, why might not a registry record link to a Medicare – a Medicare hospitalization record? Now on average I'll say that when we're linking inpatient clinical registries with Medicare claims data, we observe linkage rates usually in the mid-70s, approaching 80% in some cases, and – and lower, in the low 70s in other cases. So why – why don't some patients – or some registry hospitalizations link? Well, it could be that that registry hospitalization truly is in the Medicare claims file but there's some inconsistent coding of procedures or diagnoses and so we're not quite finding them as – as we would expect. Or, alternatively, there could be inconsistent service dates or patient information that prevents us from making that link. It could also be that the registry hospitalization truly is not in the Medicare claims data because Medicare, perhaps, isn't the primary payer, perhaps the individual is in Medicare managed care and not – therefore the hospitalization is not included in the fee-for-service files. I've listed some other reasons why that – that might be the case. Often when we're doing these linkages we find – we find that the issue really is at the site level. So there may be a Veterans Administration site, for example, included in the registry, and of course we wouldn't expect there to be linkage with the Medicare claims data within a VA hospital.

Why would you stop it at inpatient data? Once – once we make this link, we've linked a registry record with an inpatient hospitalization in the inpatient Medicare claims file, then we have an encrypted beneficiary ID that we can take, and go back to CMS and ask for permission to – to then add to our – to our data. And we can – we can request outpatient carrier-type claims and we can request Part D data for



those – for those beneficiaries included in the finder file. And – and that’s what we’ve done and that’s what we’re – what we’ll be talking a little bit about today.

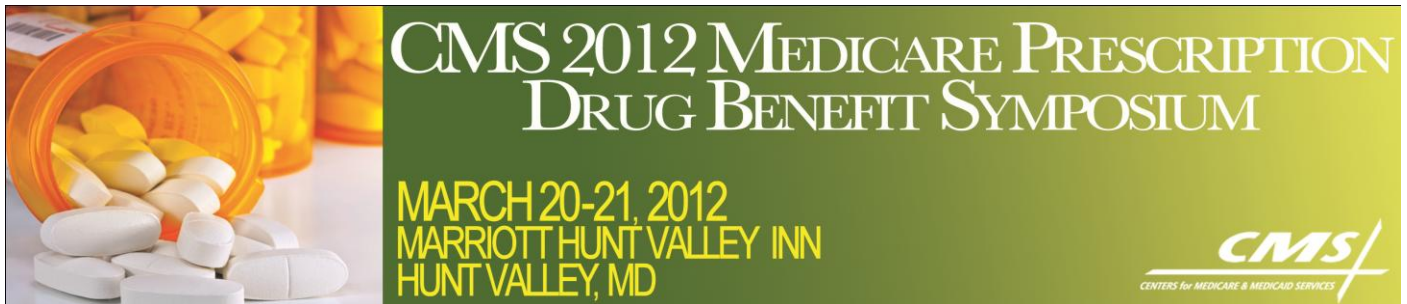
There are some real benefits of – of linkage with Part D, and some of these have been really nicely discussed already. Certainly it’s a useful platform for comparative effectiveness and safety. As I mentioned before, the registry really has rich clinical detail similar to a – to an epidemiological cohort, and Part D data, of course, brings this longitudinal medication exposure data that we wouldn’t have from – from the registry. Although maybe I didn’t say this specifically, it’s important to note that these registries, you know, collect at the time of hospitalization and there’s really no longitudinal follow up associated with the registry. They’re entered at the time they’re in the hospital and then they leave. So it’s really just that – that point in time.

Linkage of Part D data with inpatient registries can also allow us to look at medication transitions from the inpatient to the outpatient setting. This is important because there are a lot of quality measures. Were you prescribed X at discharge, for example, that we – we often use on the inpatient side. But unanswered is the question about whether or not that prescription is filled once the patient actually leaves the hospital. And then there are some wonderful opportunities for improved confounder adjustment with those Part D data.

I’d be remiss if I didn’t touch on some of the challenges of linking with Part D data, and probably the biggest one, really, is sample size. Although the clinical registries are sometimes very large, hundreds of thousands of individuals, we don’t link all of those individuals or hospitalizations to be – to be more accurate – with claims data as I mentioned before. And then once we actually identify the inclusion and exclusion criteria for a specific research question, we may be down to 10,000 or 5,000 individuals who meet the criteria. And then we think about Part D enrollment that may be at 50% or 60% in the population, and all of a sudden our hundreds of thousands may be closer to 5,000 or 4,000 or smaller. And that – that’s actually become – or that often is a real issue when using these data for comparative effectiveness type research.

The donut hole, as you might imagine, introduces some challenges as well. Because the donut hole is – has the potential for multiple exposures, that is each year an individual can face that donut hole, and we’re often interested in very long follow up over multiple years, you have a multiple exposure potential to that donut hole. It introduces some analytical complexities to be sure, and although we’re certainly learning more about the last point during this conference, we still don’t know a lot about how the donut hole effects adherence.

Now what this slide illustrates is that – that exposure to the – the Part D coverage gap, of course, can occur at any time. And these – the index hospitalization might actually be very poorly timed with respect to the coverage gap. It could occur in the middle of the coverage gap. It could occur prior to, after. And we have to again think about how we’ll handle this analytically. If the exposure to the coverage gap is prior to or concurrent with cohort entry, we tend to use just sort of a dichotomous yes/no variable to signify that yes, the individual was exposed to that coverage gap at the time. If the exposure occurs after cohort entry, then we have a different analytical challenge on our hands and may need to use some sort of a time varying covariant approach in a time-to-event model. And there may be other approaches that some of you have used. The point here is just to say that it – the coverage gap does introduce some analytical complexity into the work that – that we’re doing.



Not surprisingly data sources don't always agree. We think they should, but then they don't. So in our registries we often are recording intentions. So we record the physician's intention for post-discharge medications whereas the claims, the Part D drug event data, record the prescriptions that were actually filled. What we have found is that – of course not what we've found, what we know is that neither of these captures true exposure. But what the table on this slide shows is that these two sources don't always agree. This is actually a specific example with aldosterone antagonists, prescription of those, where we saw that, you know, 75 or 78% of the individuals who were prescribed that at discharge also had a claim of drug event data that corresponded with that. But there were 23% of – of the population for whom the physician intended for the patient to fill the prescription but there was no record of a prescription in the Part D data.

And there are things that we don't know, right? We don't know about medication exposures during institutional stays. So if we're really interested in understanding post-discharge adherence, we have to recognize that using Part D linked with – with these clinical registries, there will be gaps in what we know. In addition, we don't have any information about over-the-counter medications and those that may be filled through other programs, not – not the Part D benefit.

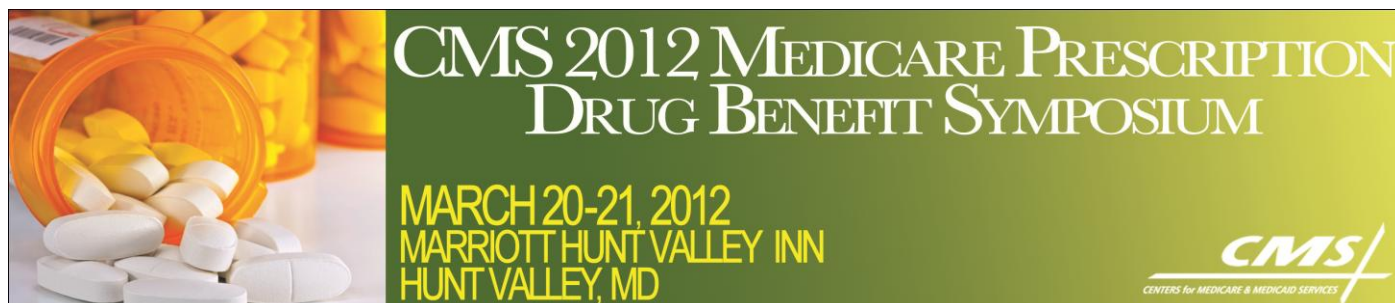
We have to be cognizant as well of the selection bias and that individuals, Medicare beneficiaries, select their Part D plan and those plans vary with respect to important characteristics. Switching also introduces the potential for selection bias. Again, we don't have solutions to all of these problems, but – but it's important that we acknowledge them and take care in our – in our analytical techniques to address them as best we can.

So to summarize, using these indirect identifiers and this approach that I've described, linkage of registries with Part D is certainly possible. And just, although I didn't note it on the slide, that approach has been published in the peer review literature, a couple of years ago in the *American Heart Journal*, and I'm – I'm happy to share that – that citation with – with all of you. There are many design and analysis issues to consider with these linked data. But if they're used thoughtfully I would argue that these data have enormous potential and would certainly be happy to – to talk with you further about that.

And now I'd like to turn it over to my colleague, Dr. Eapen.

Thank you, Leslie. So, I'm a Fellow at the Duke Clinical Research Institute, and I've been working with Dr. Curtis, and I think this study that I'll share with you is a good example of the opportunity and the potential that there is for linking Part D event data with a clinical registry to truly identify the benefits and the care of patients with a specific disease state. In this setting we looked at the differences with characteristics of Medicare beneficiaries with heart failure according to enrollment in the Medicare Part D prescription drug benefit. And this is a descriptive study to give you a sense of some of the potential once we link with a registry. I don't have any disclosures.

So just by way of background, as Dr. Curtis mentioned, the need for comparative effectiveness studies is really a priority because efficacy studies really have been stringent in terms of the inclusion and exclusion criteria that are used in clinical trials. Those clinical trial populations are very narrowly defined, and it's becoming increasingly difficult for a practitioner to be able to apply that aggregate data and the results of these large clinical trials to a local patient population. Now this is particularly true as clinical trials move off shore as well, and maybe have different systems of care, different background treatment patterns, different patient populations across the world. How do we apply those results back to our patients locally in our communities? And this is really certainly a problem in heart failure for Medicare beneficiaries. This



is not only a disease of high mortality but also of high morbidity. As you well know it's the leading cause of readmissions among Medicare beneficiaries. And as such it's responsible for 30% of annual Medicare spending.

And at the same time pharmaceutical spending represents a significant cost to Medicare as well, comprising more than a fifth of annual spending. So this is really an opportunity both to attenuate the morbidity and potential mortality of patients, but also address a significant spend for Medicare, both on the pharmaceutical side as well as the resources needed to take care of these patients with heart failure, both in the inpatient and outpatient settings.

And this has really been accelerated by the enactment of the Part D prescription drug benefit in 2006. And since then, pharmaceutical spending by Medicare has increased. And in 2008, it resulted in \$45.5 billion just for pharmacotherapies to Medicare. So with this patient population that does consume a lot of resources, faces high morbidity and mortality, what is this population of Part D enrollees with heart failure the Medicare beneficiaries? And how are their medication regimens and – and who are these patients. This has not been well described and is a ripe opportunity for future comparative effectiveness studies.

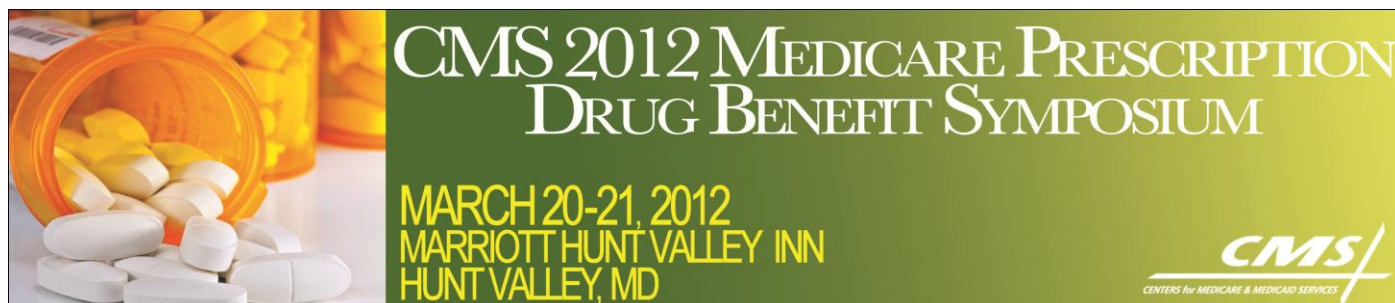
So the objectives of this study were really to compare patient-level characteristics of beneficiaries diagnosed with heart failure who are enrolled or not enrolled in standalone Medicare Part D plans. And then to describe the medications that were prescribed to these beneficiaries diagnosed with heart failure so that you could get a sense of whether they are being prescribed evidence-based guideline therapies for their care.

To do this we used the Medicare claims and the Part D event data for a nationally representative 5% sample of CMS beneficiaries, and we had data from the inception or the enactment of the Part D benefit starting in 2006 up to about 2008, and so we create two annual cohorts so we can get a sense of the prevalent heart failure on January first of each year. Our exposure of interest was enrollment in a Part D prescription medication plan, again to show what those characteristics, those demographics, co-morbidities, as well as those medications were according to enrollment in a Part D plan.

The inclusion criteria really had to do with ICD-9 codes that were relevant to heart failure. We defined our cohorts with beneficiaries that had diagnoses associated with ICD-9 codes that you see here, and these `\INAUDIBLE\` have been published by Dr. Curtis and the Archives of Internal Medicine. And we looked to see if there was a relevant CHF ICD-9 code in any single inpatient claim or at least three outpatient carrier claims during the prior year, so we had one year of exposure to really assess whether they had a diagnosis of heart failure either in the inpatient/outpatient settings. As such we required that all these beneficiaries be enrolled in Medicare fee-for-service for the entire prior year, and we also required that they be at least 66 or older, again to make sure that they had one year for us to assess whether they had a heart failure on a claim.

We then compared the demographics and co-morbidities of those patients enrolled to those that were not enrolled to see if there were any differences, and then among those that were enrolled in a Part D plan, what were the most frequent drug prescriptions, whether they were indicated, contraindicated, or maybe irrelevant to their care for heart failure.

To give you a sense first of who these beneficiaries are and what the uptake of Part D plans were, as you can see in this graph with the blue line being 2006 at the inception of Medicare Part D there was a slow uptake up to the middle of the year and then it gradually became steady. In 2007 it really was a steady



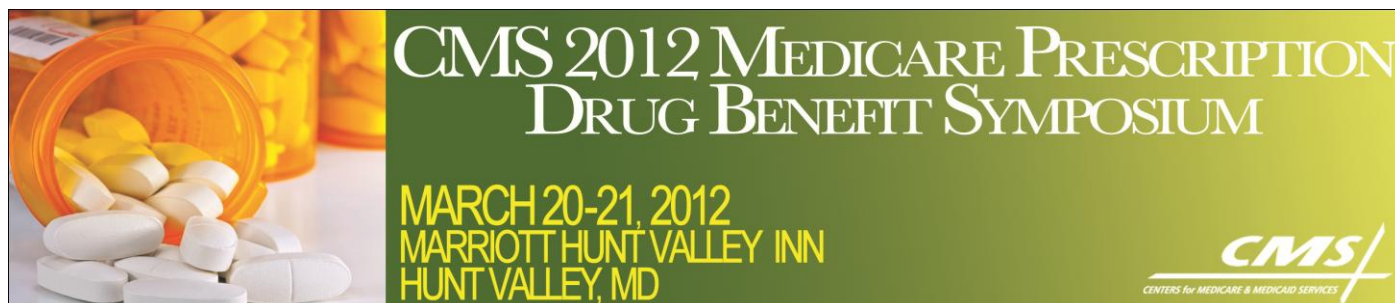
uptake of enrollment in Part D, just under 50%. So with this in mind, in the future tables I would point your eyes over to the right side of the table where we have the 2007 data, because this really is a constant enrollment pattern over that year.

And you can see that there were quite a few patients with Medicare that were enrolled in a Part D plan as well, over 50,000 patients. But you can also see that regardless of enrollment in Part D, that there was an elderly patient population here over the age of 80. And as you can imagine, this is very different from clinical trial populations. In the same trial, for example, the average age was 59.5 years. Also unlike clinical trials this was a predominantly female population. Among those in 2007 that were enrolled in Part D, only a third were male. and you can also see that there were a lot of patients who were of black race that were enrolled in Part D, more so than those that were not enrolled in Part D. You can also see that this is a lower income population than you might see in the typical clinical trial. These patients often had state buy-in status, which meant that they had some subsidies from the state for their Medicare Part A and B, so this is perhaps not only an older population, a population with – that was more representative of black people, but also a lower income population, all aspects that may not be found in a typical clinical trial which otherwise forms the evidentiary base for their care.

These patients also had multiple co-morbidities. You can see here that according to Part D enrollment status these patients had a little bit less coronary heart disease, but certainly had quite a bit of diabetes, hypertension was perhaps the most prevalent of any of the conditions. They also had chronic obstructive pulmonary disease, cerebrovascular disease, cancer, dementia, myocardial infarction, peptic ulcer disease, peripheral vascular disease. You can see how these patients have multiple co-morbidities. Sometimes this may be the exclusion criteria in a clinical trial, but in this real-world setting among Part D enrollees, we see that these are elderly patients with multiple co-morbidities and may reflect the local patient populations that we care for more accurately than some of the clinical trials.

To give you a sense of whether a heart failure patient is receiving the right medications, here you can see along the X axis a lot of different medications that may be indicated and may be evidence-based guideline-driven therapies, such as beta blocker, ACE inhibitor, angiotensin receptor blockers, aldosterone antagonists. Loop diuretics were prescribed more than a lot of these other medications, and perhaps this was mostly for congestion. And this was around the rates of 50 to 60% although it cuts off a little bit on the slide. But you can see some of those therapies that are really indicated for quality care, such as beta blockers, ACE inhibitors, and ARBs, are at a lower rate. And aldosterone antagonists, again indicated for patients with heart failure, is at a much lower rate. This was around ten percent. Some of this may be due to the fact that this data is from 2006 and 2007, and since then we've had a gradual increase in the amount of evidence for aldosterone antagonists with patients not only of severe symptomatic heart failure, that is New York Heart Association Class Three to Four. Also less symptomatic forms of heart failure. But this also may have something to do with the cohort that we're looking at. Remember, we're just defining these cohorts by prevalence \INAUDIBLE\ according to those IC-9 codes, but we don't really have an idea of what those clinical characteristics are that are associated with these patients. Are they patients with heart failure with reduced ejection fraction or are they heart failure with preserved ejection fraction? It's likely a mix of both. And as I showed you from some of those demographics and some of those co-morbidities, this is an elderly population, a very female population, a population that has a lot of hypertension. Certainly a population that is ripe for having heart failure with preserve ejection fraction. So perhaps some of these medications are not indicated in this overall cohort.

So, again, an opportunity to really link with a clinical registry, use the right clinical data that's available in a registry to understand what the effectiveness of the drug exposure is using Part D data.



And here are some of the numbers around that. Again, loop diuretics are by far probably the most prescribed with beta blockers coming up later. But again, aldosterone antagonists only about ten percent, and it would be interesting to see in more contemporary Part D event data what those uptake rates are and if they have improved.

How about the wrong medications? Here thiazolidine-diones and potentially Metformin could be contraindicated for patients with heart failure. And these medications were certainly more prescribed than aldosterone antagonists and about the same rate as angiotensin receptor blockers and maybe ACE inhibitors. So again, some medications that may not be worth prescribing or may be contraindicated in a patient population with heart failure is certainly reflective of some of their co-morbidities such as diabetes. Again, steroids, NSAIDs, contraindicated antiarrhythmics such as Sotalol, terazosin, cilostazol, all prescribed although at lower rates, and may be contraindicated. And to do this we really went through the NDC codes to lump all the medications in these different drug classes to make sure that we had a good encompassing drug class and then we looked through those NDC codes and those groups to assess the – the prescription rates.

The ascertainment window for all these drug exposures was four months after January first, remember we had a prevalent cohort as of January first. But we looked to see what the drug exposure was over an ascertainment period of four months to make sure that we captured at least one 30-day or one 90-day prescription refill in the outpatient setting.

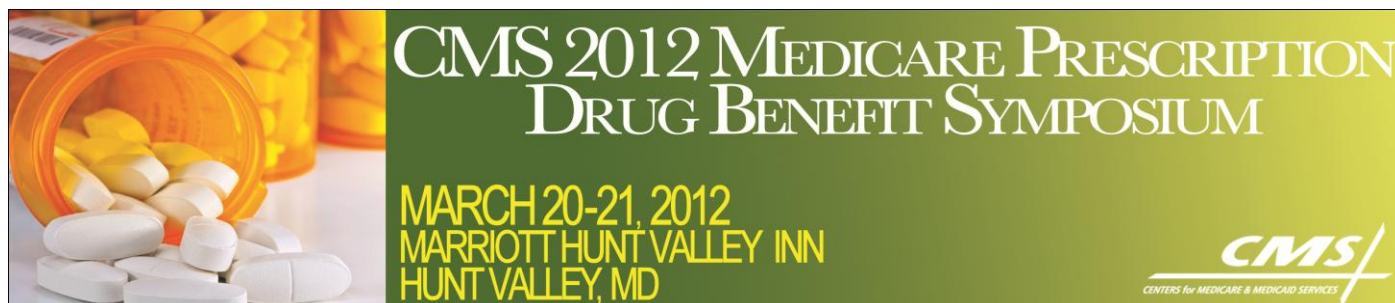
Again, diabetes medications were probably the most prescribed out of the ones that would otherwise may not be optimal for these patients with heart failure, but certainly there were other medications such as corticosteroids that also would potentially exacerbate the condition for patients with heart failure.

There's also many other drug prescriptions that were prescribed in this patient population, again reflective of the co-morbidities that you see in this real world population unlike what you might see in some of these clinical trials. Along with the high prescription rates of pota – of loop diuretics, you see that potassium chloride is prescribed pretty frequently. There's also other medications that are not related to their heart failure, such as levothyroxine for hypothyroidism which is prescribed, and again, prescribed at similar, if not better, rates than some of the evidence-based guideline therapies that we prescribe for patients with heart failure.

So in conclusion, Part D enrollees in this analysis did have multiple co-morbidities, they were more likely to be female and black, they were also likely to be of low income status as we found through the state buy-in variable, which is a potential indicator of differences between clinical trial populations and this real-world population. And we saw that these populations do differ, and so this really provides a great opportunity and ripe potential for future comparative effectiveness studies.

We see that in this patient population, again, not knowing their clinical characteristics, but that evidence-based guideline therapies for heart failure is low, but to really get more granular and to understand what the utilization rates are in appropriate patient populations, we do have a lot of potentially here to link this Part D event data in the specific cohort with a disease-specific registry, such as Get With the Guidelines Heart Failure, using the methods that Dr. Curtis has previously articulated to really get a sense of what the effectiveness of drug therapy is in real-world settings. Thank you for your attention.

It is time to conduct the assessment. Please get out your ARS response cards, and we would encourage you all to participate. As a reminder, if you are seeking CPE credit, you must respond to all assessment



and evaluation questions. After the questions and responses are read, you will have ten seconds to respond. You'll see the timer on the screen.

Assessment question number one. Inpatient registries may be linked with Medicare claims data using one, social security number, birth date and gender; two, B, hospital site and beneficiary name and address, three, C, dates of service, date of birth and hospital site, four, D, medical record number, birth date, gender and hospital site. Please vote now. You have ten seconds.

The poll is now closed. Let's look at the results. There we go. Inpatient – the correct answer is three, C, inpatient registries can be linked using the site, identifier and dates of service. I see that some of you may have chosen the first answer, which is the way that we can certainly link cohort studies where we have social security numbers, but – but typically we use the – the site and dates of service.

Assessment question number two. Inpatient registries linked with Part D prescription drug event data provide the capability to assess: time of initiation of a medication newly prescribed to dis – at discharge, post-discharge adherence to all medications, three, agreement between medication list at discharge and medications filled in outpatient settings, four, 1, A, and 3, C, or 5, all of the above. Please vote now. You have ten seconds.

The poll is now closed. Let's look at the results. So there's a – a split here between four and five. Maybe this was a trick question. The correct answer is four. I would argue that post-discharge to all medications is not possible because over-the-counter medications and those medications that are administered during – during institutional stays would not be reflected in the Part D data, but perhaps that was too – too much of a trick question. All of the above is close.