

>> So this talk is on the Medicaid Research Support Tools. And what I want to be able to do in this presentation is give you an idea. We've mentioned resources over the course of this workshop. There are also some that we haven't even touched on yet. So I want to cover those briefly for you; let you know where you can find them. Some of these are things that are available from CMS. Others come from non CMS sources. There are also a couple of new MAX Data Files available. We've been talking about the claims and enrollment data, so the traditional five files that make up the MAX Data Set. There are some supporting axillary files; I want to talk about those. And then I've called your attention to the URLs handout.

So one of the first places we began in this workshop, was the Medicaid program. And if you're new to studying Medicaid, a lot of times it's a question of how do I even read more? How do I understand what this program is about? And as Todd pointed out, with healthcare reform there is a lot changing. A lot of these resources are older, but I think to understand some of the basic things like categorically needy; what does that mean? Medically needy? To understand where the program has come from or what the foundation is, I think these are some good resources. And I'd say they're listed in increasing order of detail.

On the CMS website, they've actually created another website at Medicaid.gov that provides a lot of program information. There's information on eligibility, general information on benefits and cost sharing. There is a book that's referred to as the "Green Book". I assume that when it was in hard cover format it was green. And the version from -- that was published in 2004 has a chapter in it that does provide some more detail on the program. Over the years, the amount of information on Medicaid in that publication seemed to decrease, and by 2008 it actually had disappeared.

The Kaiser Family Foundation website is a good resource for not just Medicaid but Medicare and other health policy issues. In terms of Medicaid, they have a couple things. One is a resource book that was last updated in 2002. But again I find it's helpful for things like, well what -- just generally what does a waiver mean? Not getting into -- then there's some information about specific waivers, which of course you'd have to update. But it does provide good information about how the program works and some of the things to watch out for in terms of this idea of different eligibility across states or different benefit packages. Kaiser also publishes the Medicaid primer and the latest available there is June of 2010. Because you will be working, in most cases, with data that's historic, it can be helpful to be able to go back to those earlier editions.

Continuing on with some national resources, again the Kaiser Family Foundation, Medicaid Resource book that does have -- and then on their website there's also some reports on specific sub populations, things like the duals. If you're interested in other programs that relate to your Medicaid study, Medicare or SCHIP, you can find information there. The Urban Institute is another organization that is a good source of reports. Generally they're looking at an evaluation of an implemented program or they're looking at things in anticipation of the effect of new policy, and

it's quite a variety of issues.

The National Pharmaceutical Council, this was just a treasure trove I thought, like discovering gold, but they've closed down the goldmine. They publish a guide to state's pharmaceutical benefits that really was quite detailed. You could figure out what was included in a managed care program. If drugs were carved out or included. There were specific contacts there -- state contacts for the drug benefit. They published this from 1963 through 2005. They skipped the year 2006. They published 2007 and there hasn't been an update since. So -- but given that you may be working with some of the older data, that could still be useful to you.

And then a last national resource that I'm mentioning here is the National Academy for State Health Policy. Their topic or their focus seems to change every year. But it is a good source if you're interested in keeping up with some current initiatives or information on current program changes.

So then if we move toward state-specific information, many of you are working with one state or a smaller subset of states. One of the key things a lot of people are wondering about is waiver information. So the Medicaid.gov website, which is supported by CMS, gives an overview of different waiver types. And then in the download section there'll be reports that provide some descriptions of state-specific waivers. So you can tell something about special programs that are going on within a state. We've talked and encouraged you to consider state-specific coverage. Do they have particular eligibility criteria? So they, of course, have to provide the federally mandated level, but then they may add optional populations or services. It is, unfortunately, difficult to get at that plan information in a centralized location.

The Kaiser Family Foundation does have a database of state Medicaid plans. You can review these looking within a state. You can compare across states. The current data -- the current data that's up there is really coming from some different points in time, but they do document by footnote for a particular variable or category of care where that data -- what -- the date for that data. The source of all that are documents that the state supports to CMS, so they're just pulling that together for you.

Some of the state Medicaid agencies do actually have a website that has information about their coverage that would be helpful to you as researchers. Trying to figure out before you get data who might be in there; what's the effect of capitated managed care in that state. On the CMS website they do have enrollment statistics. Those are current to July 1st of 2010. So you don't have calendar year; you don't have a fiscal year, you just have this one point in time. They do include a trend table in that so you can look at things across 2000 and that is only current up to 2009. And then you have the state penetration rates through July of 2010.

There are the acronyms for all the various different kinds of managed care that are part of Medicaid and those definitions are given -- you'll see a download for the National Summary Report. That's where you'll find the definition for all the managed care acronyms. I think this is a helpful

resource to get descriptions of the plans.

When you look into that National Summary Report you can see enrollment by plan; by county. And that's important if you are looking at a specific part of a state or doing comparisons. You can see where is the likely effect on this managed care enrollment on the data you're going to see. Is it state-wide? Is it going to be focused on the major metropolitan area? So you can get some -- the county information I think is useful for that. And again the -- part of what you're after here is how -- what are you going to have for fee-for-service claims in the MAX data because even given the improvements in a lot of states, there can still be a significant effect on this and what you'll have for data.

And then that - there's a report from this website that also gives you information about the dual eligibles and what kind of Medicaid comprehensive managed care they might be involved in. Some further resources are descriptions of the state managed care programs so you can actually get a feel for what services are included in those specific managed care programs, what type of professionals are qualified to be considered primary care providers. People are often interested in that and that's become more of an interest as we move forward here. And then whether there were eligibility groups that were mandated to be enrolled in that comprehensive plan or if it was voluntary. So you'd see some mix of things within an eligibility group.

^M00:10:51

I wanted to just call out an example, this is older data but it -- the report continued to make it a clear way to have an example for my point. And I've mentioned that it -- in Medicaid you can't really just look at yes/no or what was the percentage in the state enrolled in managed care. So looking at this report, if we focus on the line for Georgia and we follow across, the percent in managed care is almost 96%. If you looked at this report at this highest level, I would say you'd be crazy to think you're going to be able to do a great study with Medicaid data from Georgia because 96% of the people are in managed care. As you start to look at what type of managed care were those people enrolled in? So thinking about the study you want to do; what type of managed care were people enrolled in? In Georgia, a large part of that 1.3 million were people who were enrolled in primary care case management. It's paid fee for service. No worries there, you've got the claims data. Another category for about 2,500 people were people in prepaid inpatient health plans. So that's a category of inpatient care that's being capitated. Often times this could be something like mental health. So thinking about your research question, what size population are you dealing with overall, you might need to dig into in Georgia what's in that PIHP. But the major thing is, that huge number that makes it 96% enrolled in managed care, they're in this prepaid ambulatory healthcare plan. So it's not a comprehensive plan. That's where the majority of your -- that's what's driving that 96%.

So what is that PAHP? Well in Georgia, those 1.3 million are enrolled in a capitated transportation plan. If your research question involves access to care, physical access, how -- is there any assistance to getting

to a clinic? A doctor? A hospital? Then this might be really important because you would be looking for transportation claims; things covered by Medicaid. But if you are not worried, if that's not part of your consideration then Georgia is certainly a state that you would expect to have a good rate of fee for service data; claims data that you can really work with.

Another alternative source for looking at exactly what's in the MAX data would be these data validation tables that I've mentioned. So again they are developed when a MAX file is finalized and every year of data -- there are data validation tables for every year of MAX data that's currently available to you. So they go back to 1999. You can download them from the CMS website. You'll find a lot of information on enrollment in general, but looking specifically at managed care, there are some that are just enrolled at any point in the calendar year. You'll see, and that is aggregated by eligibility group, or you can look at that based on some major categories of managed care. And then at the point -- just a point in time for June of that calendar year of data, you can see information for the ever enrolled aggregated by eligibility group. So that's another way to get at the information of how many -- who's enrolled in managed care.

For Medicaid statistics, there are lots of different resources out there. Any that I've found, trying to get at a geography smaller than a state really isn't possible. You might find things on a state Medicaid website that have information for you by county, but generally all of these statistics, if they go below national, they're just going to go to the level of state. There's information on enrollment with demographics, eligibility groups, expenditures and utilization by, generally by, those broad type of service categories, so inpatient, outpatient. You won't find statistics in these resources by diagnosis. You won't find things down to a procedure code.

A few to highlight here, the MAX chart books are available from the CMS website, they've been published three times so far. The data is from 2002, 2004 and 2008. Also from the CMS website there is the Medicare and Medicaid statistical supplement. So if you're looking for data on statistics on both programs, that's a good resource for you. The tables for Medicaid do include some trend tables that go back to 1975. So if you really want some of the history far back, that's a good source to look at trending information.

There's also information on prescription drug utilization available from the CMS website and they were published for 1999 and then 2001 through 2008 are currently available. You'll find some reporting there by broad, therapeutic classes of drugs. So you won't find things by NDC code, but it is helpful if you're wondering what -- about certain categories of drugs. And some of these tables are available in an Excel format. So if you want to work with it further or combine it with other resources you can.

From the CMS website you'll also find access to MSIS data tables. So these are federal, fiscal year 1999 through 2009. Not every state is available as you get into 2009. And it's very useful because there is an anomalies' report available to you from that website as well. And I would encourage

you to look at that as you're looking at those statistics.

And then the Kaiser Family Foundation has state Medicaid fact sheets that displays to easily do some comparisons. You can look at the data at the national level. You can look at particular states or you can create comparison tables, custom tables for yourself. Unfortunately you can't export those as Excel -- excuse me -- but you can export them and be able to download or save them, your custom tables, as PDF files.

Another resource for some statistics that you can work with and create custom tables and export those is the MSIS Data Mart tool. It's available from the CMS website. They ask you to submit an email identifying yourself, your contact information and just a sentence or two about your purpose for using that resource. You -- in the case with anything of course you want to be sure you understand the variable definitions and they do include a data dictionary, access to that by download. The tables are available; the whole Data Mart's available in Cognos Powerplay. You don't have to go get that software. When you open up the Data Mart, you're working in Cognos and I have found that Cognos, seems to work best in Internet Explorer using Firefox or Safari, you don't even often know there is a problem-- you're not getting an error message, but eventually you realize that you're getting strange results. So that is one where as a Mac user I do have to say Microsoft triumphs there. So again, be sure you look at those anomaly reports when you're using the Data Mart.

This is just a screenshot to give you an idea. When I talk about the Data Mart what does it look like when you open it? So there's certain variables, and if you click on the folder for that variable, you'll get a drop down that tells you what the details are -- excuse me -- that you can look for there. And then you'll see your results and again, you can keep working with this and drill down; you can look at things only by certain categories. It's a way that you can actually manipulate statistics and not be limited just to a published table, and that's freely available to you.

Other Medicaid data, the -- some people are interested in looking at drug utilization. Often times if you're looking at a drug that is - if you want a - is a more recent drug, before you purchase data you might want to know how is adaptation or utilization of that actually going in the program? What year of data, if the drug was introduced in 2005, what year of MAX data would I need to really see what looks like a fully implemented adapted utilization. These are downloadable data files. They are updated generally quarterly. The most current data available is for 2011 by quarter four and you will find some variation in terms of how well that quarter four for 2011 is populated. Not all states have submitted the data. This data is available as-is. There have been a couple cases where people have called and said, I'm looking at 2009, it looks wonky, can I get any updated information? And I've contacted CMS and the answer is, it's as-is. So the states submit it; CMS publishes it, if there were corrections made later, those corrections generally are not pushed forward.

These are zipped files that extract to a text format and then you can work with that text and import it into another program. They are too large to

work in Excel. I think at one point in time you could fit Delaware into Excel, but at this point I believe they're all too large -- the state files are all too large to work in Excel.

Again, it's useful for looking at some background information, frequency of use, some people have used it to calculate some power analysis. Maybe you're working on a grant proposal and you need some statistics for that and those files are useful there.

This is just a screen shot of the actual record layout for the files, so that's available as a download too. NDC information is something that, in many cases, you'll need as that's what's available in the RX file for MAX, as again just the NDC code. It's not a drug name. It's not a category of drug; you just have an NDC code.

A challenge then is finding an accessible -- and most of us would like that to be a free source -- of information in order to find out what that drug really was, or categorize drugs by a broad thing, such as antibiotics or ACE inhibitors, things like that. The FDA does have a database that's downloadable. It seems to have a smaller number of codes than the commercial products. So it may not be updated as often and they no longer support any kind of therapeutic class table. And that would be that way of aggregating multiple drug codes into one category, like an antibiotic. There's a lot of commercial products out there. If you're planning to do work by -- that really requires you to use prescription drugs as information, I'd say trying to check, if you're in an academic setting, can you partner with somebody in the pharmacy program? Trying to get access to one of the commercial databases would be a really good idea. I don't know if anyone has found any other free source.

^M00:22:57

Okay I'm going to attempt to summarize this and then we'll take the next question, but just to be sure I have it, if I missed something; make sure I get it. So the resource is the NCQA website and in there, tools for the HEDIS data, H-E-D-I-S, you may or may not be familiar with that. They do have tables for some codes that are grouping things by a drug category, so you will have the NDC code and you will have the drug name. And then an additional comment was that that same resource will have drugs by disease state and that can be helpful, looking in terms of utilization, but sometimes people are actually using prescription drug claims as a way of identifying a clinical population.

And then we have another question or comment.

^M00:23:54

[Inaudible off-microphone question]

^M00:23:56

>> Okay thank you. That comment was that he's found that the HEDIS data is missing some drugs and some of the codes won't be in there. The resource that he's used is the Lexicon database which is provided by Multum. It used to be available for free or major portions of it to academic researchers. The last time I checked, and I think we've looked at that recently, and it still was not available as a free tool, but it's another

commercial thing; I should add that to my commercial list. Okay good. I'm -- surprised that HEDIS one got away from us and it's -- I've asked this question every year for quite a while. I'm glad we finally got another resource there. I'll be sure and add it, thank you.

Okay, moving on then, some other data support tools. Of course data documentation and this is all available from the CMS website. They have a couple of specific, I'd call it like a home page almost for Medicaid data and there's a lot of links from that page. You can find -- if you're working with the data across time, I've mentioned that there have been some additions or slight changes to variables. There is a crosswalk table available so you can see what happened across time; what was added or deleted. You'll find the links to these data validation tables, which I think are helpful both pre and post acquisition of the data, or maybe you never get the data, but you're looking for information about population or the program. And then they also have, let's see, for the data validation tables, when you're downloading them you can - will also be able to download the methodology that was used to create those measures. So if you're trying to duplicate that in data that you have, or you're wondering about inclusion or exclusion, you can actually see that and understand what happened there.

Another important resource then from CMS with the MAX data is the data anomaly's database. So as they produce MAX and they note different things, either increased number of claims in a certain category, things that look like wildly increased or decreased enrollment, they will contact the states and work with the state. If they can't resolve the issue, they'll provide documentation on why that's happening. So at least explain the anomaly if it can't be resolved. There may be some times it's an issue with code mapping. It might have been things that happened with enrollment, and those will be provided. For some of the older data, the first thing you'll see, and it's a text report, the first thing you'll see is general anomalies that apply to all files and all states. And then it gets more specific to state and file information. With the more recent years -- I believe it begins with 2005, they start providing the anomaly information in a table, which is a lot easier to quickly review and work with. But I'd encourage you to look at those, especially before you request the data and in particular if you're thinking of working with something like 5, 10 states. That can help you figure out which states to work with.

We also now have some great analytic guidance reports, that were created by Buccaneer. That company has the CMS contract to do data distribution for CMS. A few that are of particular interest for Medicaid data users, there's a MAX data user guide, one on getting started with the MAX files and a very recent one on looking at what are the options across Medicare and Medicaid files for identifying dual eligibles -- and they actually have reports on, you know, what are their -- how many people do they find using one method; one variable. How does that compare to others? And then there's recommendations at the end of the report so you can get a feel for that. In addition to these on Medicaid, if you're working with Medicare with the part D Medicare data -- the drug data -- they have other guides for getting started and other reports based on that. So that's a good resource there.

If you're working with MSIS data or even working with the MAX, you can access the MSIS data dictionary. So in MAX, if the source of a variable is listed in the documentation as MSIS, if you're wondering, what was the state told? What were the rules about populating that, go back to the MSIS data dictionary and there are three attachments that I call out here that I think are particularly useful, even if you're just working with MAX; if you don't want to go back to MSIS.

And these are the eligibility crosswalk, a type of service reference table and then the program type references. You'll see a variable or value labeled foster care children, nursing facility. What does that really mean? And if you go to these tables you can see the citation for the federal code. And if you go to the federal code then you can actually see who are they talking about including in that group? What is the state mandated to include in an eligibility group? What is the federal definition of nursing facility in that -- in the case of the Medicaid program? So it can be helpful in order to have that background information.

And then a new tool that's -- a relatively new tool that's become available to us are these Medicaid issue briefs. And these were produced by actually going into the MAX data, looking at it and what's there. I had mentioned earlier the report on Encounter data. I listed out all the topics that I saw currently available and again they've looked at Encounter data now using 2008 data. It included an assessment of the OT file and I believe, I started looking at this last night. I thought it was great. I was so excited to actually have some real hard data on what it does look like. And so you're probably either thinking, I need to get a life or you've just cleared your calendar mentally for Friday night because you want to read this. But I think it's great that we really have some concrete information. I believe for 2008 they're now looking at about 27 states and they do this analysis considering the eligibility group. And I think that's really helpful because it -- if you're looking at certain groups you can see what's the effect of managed care? What states could I work with and feel that, you know, I've got good information there. With the inclusion of the OT file, again, I think that means you've got a lot of the basic claims we're going to need. They didn't look at long-term care for a lot of reasons, but I think it means that you could have more confidence in the work that you could do, including encounter records.

^M00:31:41

So now, this is another resource that's newly available. It's a file -- the acronym is ESPC, and what CMS did was have a contractor go out and gather information about the Medicaid program and the CHIP program and pull that information into a centralized resource. So it isn't the unique source, but instead of having to go off to Kaiser for one thing and another website for something else, they've made an effort to really pull all of that together. You can see from the number of variables -- almost 1,400 -- they found a lot of information. There is a data user guide, and then they do give some good examples about how to work with the data. I believe those examples are in Access. This is downloadable data, so it's -- you can have it as soon as you want to work with it. It begins with 2005 and then moves

forward. The next update will be calendar year data for 2013. I'm imagining that will be available somewhere in 2014.

There also is a file, a MAX Provider Characteristics file and this is another downloadable file. Traditionally, as we mentioned, what you had in the MAX data was a state provider ID. This file took the MAX 2009 data and started building a record using the NPI. Some of the options for variables are those -- the Legacy ID, which could be a state-specific provider ID. You've got a national taxonomy there, so in terms of specialty, you've got something that has been equated across states, and there actually are counts of beneficiaries within the MAX data for that calendar year by the provider. You can look at number of claims by that provider, and there is good documentation for this. You will note that currently it's -- there is 37 -- and I apologize, I've got two typos on this slide, but that's 37 of 51 jurisdictions. So it's not every state, but there are a number of states that are there. The issue was, how well were the states reporting the NPI.

Okay and now -- we now have some MAX files that are newly available and these are going beyond the traditional claims and enrollment, the five MAX files. These are files that are not downloadable, they would have to be part of a data request. To give a little bit of background, before I talk about the specific date of death file. In the MAX data, there is a field for date of death. And over the years, people have just concluded it's poorly reported and it comes in via MSIS and it just doesn't seem to be accurately capturing all deaths. Beginning with the MAX 2007 file, the Social Security Administration, the SSA, date of death is populated in that file and its dates of death ranging through December 30th of the calendar year for the file. So beginning with 2007 MAX, you do have good data on death within that calendar year. Of course again looking backwards into how the data got there, it's dependent on having had an SSN reported in the personal summary file. And then we now have a date of death file that's considered a current file. So this looks at people who were enrolled from 1999 through 2008. It includes a date of death as recorded in the SSA Death Master file. Any date of death, all the way through and inclusive of October 29th for 2011. So by requesting this one data file, you have any date of death for someone enrolled 1999 through 2008 all the way through October 29th of 2011. And you would link these to your MAX data via the Bene ID. So again, if you've got older data without a Bene ID, it's a matter of requesting a crosswalk file. This is not cause of death, this is only date of death.

Then a few more files that are newly available, one is called Mini-MAX. And there's been a lot of requests over the years for some kind of national sample data for Medicaid. And the general reaction was, like, oh no we couldn't possibly do that because how do you get a national sample with all this variation in there. But CMS jumped into it and this is a 5% sample. It's a cross-sectional sample that's been stratified both by state and eligibility group. There is a report that talks about how this file was developed, and I think it's really important to review that if you can. I really encourage you to review it before deciding to work with this data so you understand really how it was constructed and how that would limit, or should guide your analysis. And there is good documentation for it.

You'll find that some of the MAX variables have been removed from this sample and that was an effort to try to reduce the file size. Part of the driver for doing this national sample was to have a sample -- a national sample out there. The other part was to try to get something that might be more accessible to people if they've got hardware limitations on file sizes. So they took out some of the MAX variables. They're those variables that are constructed that seem to, like save some analytic steps for you. So an age-group variable might be removed, you're still going to have date of birth. So it doesn't limit what you can do, it only means that you're going to have a little more work because some of those constructed variables have been pulled out. Currently this is available for calendar year 2008. It is a research identifiable file. I'll be talking about that request process in the last presentation. We don't yet have the record layout that Buccaneer will be using, but you do have in the reports the record layout that was used when Mathematica created the file; they should be essentially the same.

Yes?

^M00:38:49

[Inaudible off-microphone question]

^M00:38:58

>> Okay.

^M00:38:59

[Inaudible off-microphone question]

^M00:39:05

>> Okay the question was, which resources are free or which are archived on the CMS website. Any of the reports; any of the user guides at Buccaneer, any reports on the CMS website, the state drug utilization data, anything like that is free. Anytime I've said downloadable in terms of Medicaid data, anything downloadable is free. With the Mini-MAX file and the date of death files because they're considered research identifiable files, they do have a cost involved. So there is cost for those files; anything downloadable is a free file.

Okay. And then just lastly, letting you know that other resources are available from the ResDAC website and then you're also always free to contact us via the web. You can call our toll free number or you can send an email. And we do answer the phone. Sometimes people are so startled that a human being is answering the phone. They're waiting -- they literally are waiting for the beep on the recording and they're quite surprised when that isn't it. So we do try to be there to answer the phone. And you're free to call no matter where you're at in the process.

We are -- today's an exciting day for us; we're launching a new website by the close of the day. So that new website will include a very prominent link. Contact us, submit a -- I think it's submit a question, and you'll be able to enter your question on a web form that then goes right into our ticketing system. So are there any questions about these resources?

^M00:40:55

[Silence]

^M00:40:58

>> Okay thank you.