

Medicaid Innovation Accelerator Program (IAP)
Improving the Quality of Medicaid Encounter Data
October 12, 2017

Jessie Parker (JP), GTL and Analyst on Medicaid IAP Data Analytic Team, Data and Systems Group, CMCS:

Today we'll have three speakers from different perspectives:

- Teresa Gibson, PhD, Senior Director, Health Outcomes Research, Federal Government Health and Human Services, Truven Health Analytics: She will provide a grounding in encounter data and will present examples of dashboards and other strategies some states are leveraging to improve their data quality.
- Jon Huus, Supervisor Data Quality and Analytics, Encounter Data Quality Unit, Minnesota Department of Human Services: He will provide a state perspective on encounter data and speak to Minnesota's approach to improving the quality of the data.
- Denise Love, Executive Director, National Association of Health Data Organizations (NAHDO): She will be discussing lessons learned on encounter data from state all-pair claims databases (APCDs).

I'll provide a quick overview of the Medicaid Innovation Accelerator Program (IAP) before turning it over to guest speakers. Teresa will cover challenges inherent to encounter data as well as strategies to improve qualities of data. We will then hear specific examples and processes around managed care encounter data from Minnesota followed by NAHDO's overview of encounter data and applicable lessons learned from state APCDs.

IAP is a cross-center collaboration at CMS intended to support state payment and delivery system reform efforts. We are in the data analytic functional area and one approach we are using to increase data analytic capacity within state Medicaid agencies is to host webinars such as this one on data-related topics and challenges that are relevant across a wide array of states.

Our primary goals for today's webinar are to communicate the importance of high-quality encounter data and to provide various strategies for monitoring and improving quality. We will cover data validation, quality dashboards, data cleaning, benchmarking, and other best practices. The hope is you will walk away with specific approaches you can apply in your own local data environment.

Our first speaker is Teresa Gibson. She will provide an introduction to encounter data and background on recent trends in Medicaid managed care as well as several approaches to monitoring and cleaning encounter data.

Teresa Gibson: We'll start with a definition of encounter records. This slide contains some information on encounter data records including patient information, provider information, diagnosis codes, service dates, place of service, and procedure codes. Encounter data, like fee for service claims, allows the health plans or managed care organizations (MCOs) to provide information about medical services to the state's data repository. An encounter record can contain no or partial transactions and information about those financial transactions. This is due to payment arrangements such as capitations.

States have been shifting the administration of their Medicaid benefits to MCOs to improve access, improve quality of care and create more stable funding streams and improve value. In doing so, the MCOs

take on financial risk, through capitation, which is either full or partial. In a full risk or fully capitated managed care arrangement, the state pays the MCOs a capitated per member per month (PMPM) rate, and the MCO assumes the risk of the plan cost. The impact of this shift is that it reduces the state's visibility into claims data since the data is now housed by the MCO, so all services no longer generate a separate financial transaction.

Here are some 2014 statistics:

- 77% of Medicaid beneficiaries were enrolled in MCOs. Managed care is not a main delivery system in over half of the states.
- 39% of all Medicaid dollars were paid to MCOs.
- There are 600+ comprehensive Medicaid MCO contracts.

For states with comprehensive managed care programs, managed care penetration varies from less than 1% in Alabama, Arkansas, Idaho and North Carolina to nearly 100% in Hawaii, Puerto Rico, Tennessee and Washington. Each state program design is different and so are the populations they have chosen to enroll in managed care.

So high-quality encounter data is needed for many reasons:

- Risk adjustment
- Program oversight and integrity
- Quality measurement

For example, risk adjustment requires an accurate and complete risk profile gathered from the data elements that are submitted such as complete demographic information, clinical information, diagnosis codes, and at times procedure codes. So there are many good reasons that high-quality data is needed.

There are challenges to high-quality encounter data:

- File formats: There are typically multiple versions of file formats for encounter data submissions, and MCOs must modify the reporting to align with the state-specific submission requirements.
- Rejections for a number of reasons: Non-state claim formats, claim adjudication edit, procedures not covered, retroactive adjustments, out-of-state providers, issues with provider IDs, unregistered taxonomy codes, and encounters missing a provider.
- Variations in timing and quality for completeness of data collection and transmission for providers of care.
- Coding and completeness can vary greatly. However, the introduction of ID10 in 2015 may help to improve granularity of information.

In terms of encounter data cleaning, states may be able to develop front end edits to better accommodate encounters, recognizing that some fee for service edits may not be appropriate for encounters. In terms of modernization, states should continue to invest in updating systems to reflect the evolution of claims and encounter data. States can develop dashboard summary reporting requirements to enhance oversight of changes in encounter data quality. Finally, states and MCOs should work together to educate providers and develop processes to simplify management of provider rosters.

We have been looking at quite a few dashboards and examples and have brought them into this presentation. So states can develop dashboard summary reporting requirements to enhance oversight of changes in encounter data quality. This is an example we're showing you from California of the average lag time and days by submission date. So one could create something like a quarterly performance dashboard and it gives a comprehensive view of managed care performance across the plans. It covers enrollment by beneficiary type perhaps, eligible demographics, utilization and grievances, and it could use encounter data to calculate the utilization component. This is just one example.

The second example is Washington State and this dashboard provides an overview of encounter data quality as well as HEDIS and CAPS measures by plan. Tennessee does something similar. This uses a checklist to check the frequency and time of submission of data by plan and it uses define thresholds and goals for plan data. We see the circles showing a percentage here.

This example is from Florida. Florida, California and Missouri hold regular webinars and conference calls with their MCOs, and they can talk about best practices and answer questions, address encounter data problems and talk about submission requirements and other guidance.

Tools are very important in maintaining and improving encounter data quality. These can include data scrubbing or data auditing programs, including those that have data anomalies and correct them, and those can have a high payoff. These can address things like inconsistent field lengths, inconsistent descriptions, inconsistent value assignments, missing entries, and violation of integrity constraints. Optional fields in data entry forms need particular attention and can be a significant source of inconsistent data. So limiting the use of these optional fields, providing guidance for populating optional fields, and paying particular attention to optional fields is very important in maintaining data quality.

Here's another example from New York State that provides a good example of specifications and data anomalies that will be flagged and sent back to the plans.

To improve data quality:

- Hold regular informational encounter data webinars with all plans, reviewing best practices and answering questions. California and Missouri do this.
- Develop clear submission requirements, definitions, and data specifications in the managed care contracts.
- Implement automated data validation and reconciliation processes. Kentucky has examples of encounter data submission requirements. We found that in Washington State they set up a single support for collaboration and information exchange, which also provides dashboards. In New York State we have found requirements for submission of data on a monthly basis representing all covered services in the contract, and this includes new records, adjustments and prior reject plans.
- Setting consequences for performance. In California, they set consequences for different levels of performance, and consistently underperforming plans are subject to penalties. As an incentive, when Medi-Cal auto-assigns enrollees two plans, plans with higher encounter data quality get a higher allocation of enrollees as an incentive. Also encounter data is a regulatory requirement in the managed care regulations, and this is specified under 42 CFR 438242, and this speaks to plan data to states, and there's a separate regulation regarding state data to CMS. So these

requirements for encounter data should be stipulated in the MCO contracts. States use penalties as well as incentives for compliance with this element.

To evaluate encounter data quality, first compare benchmarks. So comparative fee for service data quality, compare by plan, compare over time. Kentucky has a Medicaid encounter data benchmarking study that can be found online. New York has validation reports for encounter data. In New York, PMPM cost data comparisons are made between health plans' submitted costs and reported costs. So then health plan results are compared to determine which plans had sufficient reporting to be included in the development of risk adjustment models and payment rate development. Washington and Tennessee do something similar. They use checklists to check the frequency and time of submission of data by plans. They use defined thresholds and goals for plan data, and they also look at operational metrics such as member wait times and grievances. These reports monitor compliance against contracts. Another example is California, and they measured data quality based on completeness, accuracy, reasonability, and timeliness. Examples of measures California uses are: the percent of duplicate submissions as a completeness measure, denied encounter turnaround time, a reasonability measure, within a certain amount of days; average lag time by service data for timeliness; and they also track usage of key fields like diagnosis and revenue codes.

Finally, it's important to standardize file formats and submission processes for all health plans.

Next we have an example of a benchmark, of reports. This shows HEDIS versus encounter rates. It shows accepted encounters versus 2014 HEDIS rates. This can be a great visual display of that type of comparison.

Finally, here's an example of a data validation report. It's a New York MCO example by field, and this can be found online. We're looking at the percent number of missing valid data. This is also a great way to look quickly across fields and see if the data are populated and complete.

Jessie Parker: Thank you. Now for questions. I'll turn it over to Tracy Yee from Truven Health.

Tracy Yee: Our first question is: *What is the impact of HIPAA on encounter reporting?*

Teresa Gibson: I'm not a HIPAA expert, but I know that encounter reporting follows a HIPAA format similar to the 8037 healthcare claims transactions, and the encounter record layout the MCOs or agencies receive is typically compliant under HIPAA guidelines. In the future there may be changes that have a direct impact on the way MCOs and agencies capture and submit data, so states should continue to monitor HIPAA regulations as they become available.

Tracy Yee: *You discussed quality benchmarks used for encounter data. Can you talk more about utilization benchmarks that could be helpful?*

Teresa Gibson: Yes, and my experience in this comes from the Medicaid encounter data toolkit, which I see someone in the chat box had mentioned that, and this is an important resource, also. There are several sources of comparison data we've been looking at for benchmarking and reporting. The first would be fee for service data. So states transitioning, this may be the primary or really only source of comparison within the state, so it's a good place to start. Looking between if there's still fee for service plans, comparing those to the managed care plans, that would be a meaningful comparison, and then fee for service data also perhaps from other states. Historical data also would be important trend information, so looking at

data over time. This can be important also for data quality checks to see if fields are coming up populated within control limits. Looking at quarterly, annual and month-to-month data. So time is a very important aspect of monitoring data quality and examining completeness. You can even apply control limits to the time trends.

Another thing would be just to take a mean across the plans and compare information between the plan and the overall mean. This could be a useful piece of information. You could also turn it into a benchmark by looking at the distribution and say the 90th percentile or something of all the plans and create more of a threshold benchmark. In thinking about comparisons across plans, though, I would think that risk adjustment is important, demographic adjustment, so some plans may be very different at compositions that would be good to risk adjust or to do some demographic assessment. There are some national data sources to establish benchmarks, so the MSIS Datamart. There's other sources of aggregate national data, the Kaiser Family Foundation reports. These comparisons can be made to overall national sources.

Finally, plans' financial reports, these can be used in ways to see what data is being received in matches what was reported in the financial report, so it can be a good check or quality, sort of like a check digit, almost. But in terms of are the rates being reported on certain measures, can you actually line them up with what's required in the financial report? That's also a good source of benchmark information.

Tracy Yee: *Do you have any insight into how states are using CAHPS in relation to encounter data?*

Teresa Gibson: It's important to link CAHPS to the encounter data if it's reported to look at satisfaction measures. So I think it's an important link to make. It is at the plan level, so it's something that can be an indication of many operational issues and also just overall satisfaction and the ability to link, too. So it's an important link to make and one that operationally wouldn't be too difficult to make.

Tracy Yee: *Have you found that any states are conducting quality services beyond formatting or valid values? Are we seeing states do anything beyond standard formatting quality comparisons?*

Teresa Gibson: Sure, you can put in many hard edits just in terms of basic demographic fields but also in terms of looking at the concordance between gender and a diagnosis code. It can go pretty far. I come from the private side of the world and we see pretty extensive editing and if it's not there yet it will move in that direction.

Tracy Yee: *Is Arizona Medicaid doing anything similar to what we've talked about?*

Teresa Gibson: I personally don't know but I'm sure it's something we can look into. Tracy?

Tracy Yee: I'm not sure. I would imagine they are because they're such a heavily managed care state, but off the top of my head, I'm not aware of what exactly they're doing.

Another question: *What do you think is the most important strategy for encounter data reporting?*

Teresa Gibson: The most important strategy is to set expectations with the plans and then start to monitor the reporting levels and then give them clear feedback on a regular basis. I think that that would be the core of a reporting strategy that could start with maybe a limited set or a limited focus and then expand from there.

Tracy Yee: *Are there data areas that are the most important to focus on for high-quality encounter data?*

Teresa Gibson: There's two areas to focus on as a priority. The first would be just the reporting percentage. Are all encounters being transferred from the MCO to the data repository? That can be kind of difficult to determine, especially when there's not a financial transaction always occurring, but understanding whether you have full and complete data or not is just definitely the first place to start. Then complete coding of diagnosis and procedure codes as well as basic demographic data, because a lot of the quality and risk adjustment algorithms operate off of demographic data and diagnosis and procedure codes.

Tracy Yee: *Will we have samples of the data evaluation templates available?* We will be posting the slides on the website. Back to Jessie.

Jessie Parker: We'll shift gears slightly to hear from a state perspective. Jon Huus from Minnesota will talk about their approach to ensuring accurate and timely, complete and standardized encounter data.

Jon Huus, Supervisor Data Quality and Analytics, Encounter Data Quality Unit, Minnesota Department of Human Services: We're the Medicaid administrators for Minnesota. I have a ton of stuff so will move quickly. I could talk for an hour and a half on any one of these slides. Anyone who wants to contact me, my email is at the end of this presentation.

The agenda is:

- Mission: Accuracy, Completeness, Timeliness, and Data Standardization
- About Minnesota Medicaid
- Encounter Claim Data Process Flows
- Where Managed Care Data Often Becomes Compromised
- Strategies
- Lessons Learned

About Minnesota Medicaid, we have a population of about 1.2 million enrollees at any time, probably a little closer to 1.4 million now. Around 75% of our enrollees at this time are prepaid managed care and 25% we handle here in a fee for service arrangement where the fee for service plans are paid by us here at DHS. They have a full claims shop. We have currently eight Medicaid MCOs and we expect that number to grow pretty dramatically the next year or two. We've had as many as 12 at a given point. We're down to eight because a couple of them were consolidated, but with the advent of more ACO activity in Minnesota, alternative payment methods, we're going to see I think many, many more MCOs.

Right now the unit I supervise we have seven people in my unit. It's completely dedicated to the quality of managed care encounter data and over the last three years we've shifted from a mainframe focus to a data analytics focus. It's been a little painful in some ways and incredibly beneficial in others. For one, we've had to deal with getting appropriate staff into positions that can do complex data analytics through SAS and our data warehouse. The focus of quality care in Minnesota DHS is through analytics. We have the SAS platform. We have a Teradata data warehouse. We have automated web reporting environment for feedback to the MCOs. We rely heavily on MMIS capabilities for gathering the data and editing it front end before it's migrated to our data warehouse. The unit I'm in is closely associated with the healthcare data analytics groups within DHS. I'm really proud to say that Minnesota has a very powerful research and analytics area. Currently we have quarterly meetings with data analysts within DHS and quarterly meetings with our MCOs.

Again, our mission is these four things: Accuracy, Completeness, Timeliness, and Standardization. Standardization for us means that the health plans have to submit data to us through our MMIS claims systems on X12 formatted claims: NCPDP forms, 837i's, p's and d's. What that does for us is give us two powerful advantages. One is that the data gathered can sit right next to our fee for service data in our data warehouse and it's available for analytics right next to our fee for service data, so that's a really great thing. The other great thing about bringing data in through MMIS is that these claim formats are HIPAA-regulated. HIPAA technically doesn't apply to gathering managed care data through MMIS systems. HIPAA is more of a mandate from provider to payer transactions. This being a payer to purchaser transaction is a little different level. We have some leeway with HIPAA so we can adjust what our expectations are with the health plans as far as the managed care data goes.

The other advantage to gathering the data through MMIS is that it allows us to do really extensive editing of the claims on their way through the MMIS system before they hit the data warehouse. We don't change the data on the way in. We add fields to it like things like calculated paid amounts and things like that, but the raw data as it comes in remains intact all the way to our data warehouse, and that allows for really powerful analytics on data integrity.

Accuracy speaks for itself. We want the data to reflect accurately the services our clients have gotten.

Completeness, we want to make sure we get all of the claims. This is probably the toughest thing to measure and we'll talk more about it.

Timeliness, we want the encounter data from the health plans as quickly as we can.

I'm going to go briefly through our claims process flow. From the time client receives service, the provider bills MCO by claim submission, the MCO adjudicates and "pays" provider (some payments are \$0). MCO claims are moved into varying types of data warehouses at the MCOs, which is one of the things we deal with here. A lot of people think the claim comes directly to us intact as it has been received from the provider but that's not really true. The claims go into the MCOs, into their warehouses. The MCOs then regather the claim from those data warehouses, creates X12 and NCPDP encounter claim files submitted to us that then go through our MMIS system, which is a parallel to our fee for service claims. Claims data is stored in the mainframe system.

From there, we do all the analytics on the data in my team to make sure does the data have validity and integrity. Then our researchers within DHS do their analytics from the same data source.

This slide I use for training. There's been historically a lot of questions for probably me in particular where people say, "Why isn't it that the health plans don't just submit us the claim that they got from the provider?" In the upper right, this is the fee for service flow, the little blue box on the upper right. The claim goes from the provider to DHS or the payer and then we pay it. It's very straightforward. If the data's not solid and valid, then the provider doesn't get paid, so that tends to be a very nice process for data submission because if it's not right then the doctor doesn't get paid or the hospital or whatever. Down below on the managed care flow you can see it's a little more complex. The provider submits claims to the MCOs, the MCOs run the claims through their claims system, and then into their databases. In the square box there it says programmers and analysts like us reassemble the encounter data and produce a claim form in X12 or NCPDP format and then send it on to us.

First you can see that the payment incentive isn't there anymore. The health plans have been prepaid and so the incentive to get us the good data is a step removed. So we believe that the regathering of the data from the data warehouses at the MCOs, and where they regather that and put it into a claims form to submit to us is where the breakdown of the data integrity occurs.

I'm going to go over briefly some of the strategy we use in Minnesota to control the quality of encounter data. One is we believe that extensive, timely feedback to the MCOs on the kind of errors or problems we're seeing with the data is incredibly important. Probably the most important thing we can do is to give feedback. How can we expect good data from the MCOs unless we tell them what kind of problems there are? It's for a number of reasons. It's easier for us internally here to identify problems at times with the data we get than it is for them. Their data comes into our black box. We get the edits. We get the kickouts. We get the raw data to work with. They don't get as much.

To talk a little bit on reporting, these reports on your screen may be difficult to read. The slides are available, but briefly these are a sample of some kinds of reporting we have been using to feed back to the MCOs. On the far left is a claims life cycle report that shows how long it takes for claims to get to us. Each horizontal bar there's four segments: 1) Date of service to date health plan gets it; 2) How long from they get it to time they adjudicate it; 3) How long from adjudication until it gets to us at DHS; and 4) For us internally at DHS how long does it take to make it available for analytics in our data warehouses.

Each one of those horizontal bars is an MCO. This is actual data. Each MCO varies quite a bit. The segments within those bars also vary quite a bit. This is pretty new, this version of this timeliness reporting, but we think it's really informative.

Next, the second to the left where you see the redacted section is our daily edit report. This report is available to each of our health plans in a web environment every day for the claims they submitted the day prior. It shows them the health plan's claim number and what kind of editing if any occurred on it. One feedback we've gotten from health plans is they want to know right away what happened with their claims rather than waiting for two weeks to find out on the remittance advices or in our other files. Third from the left is the biweekly failed voids. The void process in our system is very important because it's a way for the health plans to say oh, here's a claim we shouldn't have sent you before. It gives the health plan a way to void the claim out of our system, and if those voids fail then, then we're stuck with data in our data warehouse that shouldn't be there. So failed voids are really important.

The chart on the far right is an example of another report, a percentage of the health plan claims submitted that were duplicates. We don't like duplicates in our system. Identifying duplicates is very difficult when you get right down into the nuts and bolts of it. It's actually an art more than a science, but you can imagine how it messes up things like rate-setting and analytics and HEDIS and risk adjustment and all those things if the person's service is counted more than once. So we give the health plans a lot of feedback about what percentage of their claims are duplicates and which claims were tagged as duplicative. This is some reporting we do to fulfill the strategy for controlling the quality of our data through transparency.

The corrective claims penalty is another strategy we use. Without going into great detail, my staff has developed a process by which health plans are accountable for correcting claims that come in that have errors. There's a period of time given for each error to be fixed and if the error isn't fixed there's actually a fine for that line that has the error or errors on it. The fines right now are pretty low, it's \$1 per line, and

a given error a fine can be imposed up to four times for four quarters. If the error's corrected, then of course it goes away and they're not fined.

The overhead for building and maintaining this has just been huge. It was very, very difficult SAS programming and it's tough to administer, but I think the health plans have really responded extremely well to it. We see a lot of activity of the health plans and in dealing with the errors we tag through this program so that they come through and make corrections, we have eight really excellent health plans and they have really responded to this. I imagine myself sitting in a health plan as a staff person having to go to my manager and saying "We have to cough up this amount of money in penalties for errors." It does raise eyebrows I'm sure at the health plans. The intent is really on our end to try to get erroneous data corrected. We believe this really does raise a lot of attention to that. This all is going to be errors, and this is just one of the ways we try to encourage the corrections.

So benchmarking is another strategy. It's our way really to assess completeness of data submissions. Has DHS received all the managed care encounters we expect to receive? The process is that we compare the expected claims submissions and paid amounts to what we actually get. The thing that's really tough about this one is creating the expectations. It's very easy for us to count what we get. It's very easy for us to sum up the dollars that have come in as paid amounts on these claims. What's really, really hard is to create meaningful and reasonable expectations. The reason it's so challenging is because we've had the Affordable Care Act coming in that has caused a massive redistribution of the health levels within our population, therefore creating differences in how many claims we get, and also a redistribution of people among our health plans has caused the mixes of healthy versus sick people within the MCOs to really be a tumultuous change. So creating expectations for plans where their healthy-sick mix is changing so much is really, really difficult. We're working on risk adjusting our expectations of volumes and dollars. Right now frankly with sort of mixed success. We use ACGs at this time. We're always looking for better ways to build expectations of how many dollars and how many claims we should be getting from each of our MCOs.

Assurance protocols is another strategy. This was legislated by the Minnesota Legislature about three years ago where they wrote a statute that told Minnesota DHS that we had to go out and hire a third-party vendor to help us build a set of 10 at protocols we needed to implement to ensure the quality of encounter data in Minnesota. The quality of encounter data in Minnesota is really, really important to the Senate Health Committee and so forth. We hired the vendor and together with them we came up with these 10 protocols (from slide):

- 1: Timeliness of Submissions
- 2: Resubmissions
- 3: MCO Quality Checks Against Benchmarks
- 4: Duplicate Encounter Records Submitted
- 5: Rejections and Denials by DHS
- 6: Control Reporting and Reconciliation
- 7: Claim Reviews
- 8: Remediation Plans
- 9: Data Quality Assurance Report
- 10: MCO Review of Provider Data

There's some mixture of these protocols in the strategies we have and will talk about. For instance, the duplicate encounter records; timeliness. These protocols in their implementation created some of the other strategies that we built. This has been a really good thing in that it's written into the contract that the health plans will comply with these 10 protocols into the MCO contract, that they'll comply with these. Given that it's in the contract, the health plans do pay attention and we have really good cooperation with these so far.

Some of these 10 are a lot more robust and meaningful than others. The Legislature in that same statute told us we had to hire an independent vendor to come and assess how well we implemented these quality assurance protocols, and we've done so. That assessment was completed about three months ago and we were given good marks on how we implemented it and given some suggestions but overall it's gone really well. The report on that implementation still has to go out to our Legislature.

Another strategy for controlling quality is data editing. Again we have implemented edits in our MMIS claims system that managed care encounter data is subject to. When my unit became an entity, we wiped out all of the fee for service edits that were applied to managed care. We just turned them all off or turned them to warnings because the managed care business isn't necessarily subject to the kind of policy constraints that our fee for service business is. The classic example is that our fee for service system will not pay for a physician to perform a chiropractic service, but in the managed care world, we offer the health plans the freedom to implement those sorts of policies that they see fit within their organizations, so we don't edit that like we would in fee for service.

So we ended up turning all the fee for service edits and created all the edits you see here. There's three sections here. The upper right are the header edits, the middle section are line edits and the lower section are warning edits. The upper two are what we call true denial, the header and line edits. Those edits will cause a claim to be denied in our system and those claims are not used for analytics or rate setting, risk adjustment, HEDIS, nothing. Those are the edits we like to focus on the most but warning edits are also really important because many of those give us information about a kind that are really important, too.

The last strategy is control reporting and this is the one we've had in place probably for the longest time. Control reporting is a reconciliation process where the health plans are required to send DHS aggregate financial data and we compare that aggregate financial data to what the managed care encounter claims roll up to. In other words, if they tell us they had a certain number of millions paid up physician services for a certain quarter for a certain program, we'll roll up those same encounters to see how close they match. At the aggregate level we work with them to reconcile the differences. In addition, the whole idea of transparency here, we've also given the health plans eyes into our databases as to the status of all their records. Instead of letting them actually have access to our database, we send them the data. Every claim they send us, every line, we send right back to them in massive files. Each line on every claim is a status of whether that data is usable in our database. If it's usable it's fine. If it's not usable we tell them exactly why. We tell them what the edit was and whether it's been voided or if it failed for some other reason. So the health plan can take that detail and run it against their detail and see where the discrepancies are. This has been a really good thing. I think all the health plans really appreciate getting this data back in the form of this detail.

I want to reiterate that to do the things we do, and I realize that's not the be-all and end-all, but to do things the way we do requires the ability to do rather sophisticated data analytics. We chose SAS. You could use R or some other tool that requires an extremely good database that includes not only claims

but also your enrollment files and your provider files and as many reference files as you can squeeze in there too. Editing reporting is really important. When we edit a claim, that turns a claim into an unusable claim. That doesn't really help anybody to have an unusable claim, so we're really conservative about which claims we have. We want to make sure the data is here and usable for reporting, so instead of putting edits on all the things we're concerned about we do a lot of reporting instead. And reporting is much more powerful than a binary edit in a lot of cases.

Some other issues we deal with, because of TMSIS we're now required to pull in MCO-denied claims. Recently in the last three years we started collecting third-party liability. The last point is how good is the Minnesota encounter data now? Everybody here at DHS would agree that now it is vastly better than it used to be but how do you attribute that to specific causes? Is it because of the work my unit's done? I'd like to think so. Is it because health plans have become more attentive? That's certainly true. Or other reasons? We're seeing data get better. It's really hard to attribute that improvement to a specific activity. I'm available for questions offline. This is my email: Jon.B.Huus@state.mn.us.

Jessie Parker: Thank you. I think that overview of a grounded real-life approach to ensuring high-quality data is incredibly valuable to other states facing the same challenges.

Tracy Yee: We have some questions. *Can you talk to Minnesota's process for submitting your encounter data to the new TMSIS system and if you have advice for other states trying to do the same thing?*

Jon Huus: TMSIS has been a huge multi-year project here. It's not part of our analytics environment here. It could be and maybe even should be. I'm not sure. The manager of our research area is sitting next to me probably cringing, but it's essentially a data dump that goes to CMS quarterly. It's massive and complex and my advice: I would consider handling it through analytics and a data warehouse. At DHS right now it's handled by COBO programmers through our MMIS system. They built it and I think it's good and almost there. Just as an analytics person I can't really understand why it wouldn't be in a data warehouse or SAS environment. Other than that I don't have a lot of advice. If you're not already you will need to be collecting denied claims from your MCOs.

Tracy Yee: *Do you require that MCOs send you their claim submission dates that come from the providers?*

Jon Huus: We started collecting the MCO receive date about a year and a half ago. So yes, absolutely, and it's really important, because we wouldn't be able to do those life cycle charts without them. Let me clarify the question: Do we know the date the health plan got the claim from the provider and the answer is yes. It's essential for us because we need to have a handle here at DHS on how long the activity is at the MCO. From the point the health plan gets it how long does it take them to adjudicate it and from there how long to get it to us?

Tracy Yee: *What's the volume of managed care claims you're dealing with?*

Jon Huus: About 70 million a year across the board. I don't have that number at the tip of my tongue. It's a bunch.

Tracy Yee: *Do you get both paid and denied encounters from MCOs right now?*

Jon Huus: Yes. We started collecting the claims that the MCO denied from the provider in December 2015. That was a huge project, which seems really like it wouldn't be difficult on the surface of it. It's like we collect paid claims; why not collect denied ones as well? First, denied claims are garbage by definition,

and then we have claims that are mixed. We have claims that might have three paid lives and two denied lives. We might have plans all denied or all paid, so we actually had to create a parallel process in our MMIS claims system to handle these. So for us this was not an easy process at all. Thousands of man-hours and woman-hours, too.

Tracy Yee: Do you currently maintain a list of minimum data elements either required or conditional for submitting emergency department records for the MCOs? How do you manage variations by plan type for each set of records?

Jon Huus: Our MCO contracts are virtually identical across all our MCOs now and so we don't have to deal with variations in the contract itself. We really appreciate our contract here. The contract itself is an essential piece of managed care and really enforced the fact that good data is important to us. We have what we call a companion guide that we put out online and it's available in a public area for our MCOs. It talks about all the elements we want to see within the X12 formatted claims, everything from service day, diagnosis and all the regular stuff you see in claims all the way to the extra stuff we ask for like MCO paid date, receive dates, etc. within a claim. In managed care we feel we can go beyond and actually reuse some unused fields on X12's to collect data that HIPAA doesn't require.

Tracy Yee: Are you using X12 277 to provide feedback to submitters of data?

Jon Huus: We use what's called an 835E and I'm not sure that's the same thing you're talking about. Yes, it is an X12 format and remittance advice. Yes. We also offer a lot more than that in the way of web environment reporting and our control for detail I talked about briefly. So the health plans get a lot of feedback on their plans above and beyond our X12 format and remittance advice.

Tracy Yee: How are you handling the duplicate edits against managed care plans when comparing to fee for service?

Jon Huus: Right now our MMIS system dumps all the managed fee for service care into one big giant bucket. If a managed care claim is the same as a previously received fee for service claim, the duplicate edit will set and that can set on a line level or a header for inpatient and pharmacy. So yes, the fee for service edit can happen between a managed care and fee for service and vice-versa.

Tracy Yee: What do you do with denied claims you receive?

Jon Huus: Denied claims remain in our database. Our database is complete with all claims that come in unchanged. We add some fields but they don't go away, and claims that are denied versus claims that are approved all sit right next to each other. They're just tagged as denied.

Tracy Yee: You mentioned your team has a staff of seven. How many teammates are on the health data analytics team you collaborate with?

Jon Huus: Twenty.

Tracy Yee: A mix of analysts, researchers, IT support?

Jon Huus: Right. On the research side they have academic researchers and data analysts and quality people who work on quality measures.

Tracy Yee: Do you validate any data against the provider's records such a hospital records?

Jon Huus: This is another subject I could easily talk for an hour about. The relationship is obviously a 3-way relationship. Right now we've done what we feel is a pretty good job of managing the data relationship with the MCOs. We have more to do but we've picked all the low-hanging fruit. The next steps will be actually to tie in the providers. We've done some audits. There's some audits planned, but what I'd really like to see here done is some sort of activity outside of an audit realm. Everybody hates the audits, especially the providers. What I'd really like to see is providers who have a stake in the quality of our data and partner with them to start sharing data, start sharing some data analytic expertise, to make sure that the encounter data we have here today reflects what they have in their medical records, in their billing systems. We don't have a very good grasp on that relationship yet, us with the providers.

Tracy Yee: Sure, that's difficult. That concludes the questions. Back to Jessie.

Jessie Parker: Our final speaker will be Denise Love from NADHO. Denise will talk about encounter data in the context of state all-payer claims databases and will discuss best practices for data quality.

Denise Love: I'm so glad to be here today because Medicaid is really an important data set for states and for all-payer claims databases. I hope the takeaway from my presentation, if you don't remember a whole lot, is to encourage collaboration between your Medicaid programs and the data agencies operating or planning to operate all-payer claims databases in your state. You both have a lot to offer each other to leverage not only resources to do your common work but shared analytics and just knowledge.

Today I'll provide a high level overview of APCD. First is the big picture. State data systems are driving some of the transformation payment reform in evaluation in addition to Medicaid. So state APCD, we have 16 in full production or implementation with additional coming online. Medicaid and claims eligibility are really important components of most of these systems. The use of shared public data, and I'll talk a little bit about how APCD differs a little bit from the Medicaid data, but when the data is shared publicly with multiple stakeholders it really demands that there is credible underlying data for broad buy-in of results, especially when we get to comparing practice and provider level performance. The data soundness is incredibly important.

In most systems, APCDs, the state is involved in data collection analytics and use. Another shared public data system requires provider and payer validation before that data goes public. I'm talking about APCDs. These states are collecting medical, dental, pharmacy and eligibility files including provider files from public and commercial payers. The dark-colored states are in production. They have full implementation. The darker color blues are about to go live and I think in three states they have so we have three more dark blues than this map indicates.

States vary in governance and funding levels but all of them are collecting claims from the payers in the states. The very light blues depict voluntary initiatives. Those states are not using state mandate to collect data but getting data from some payers voluntarily. What is the difference? One of the difference is these states taking in the APCDs are dealing with a huge number of platforms. It is not unusual for a state to have 70 or 80 more platforms to feed into the state-wide APCD if you count carve-outs, mergers, acquisitions, pharmacy benefit manager data. So they are dealing with a massive amount of platforms. We talked about volume. One state last week said they have 1.8 billion medical claim lines and 930 million pharmacy claim lines. I don't know if that's just from one year, but it is just a clue to the volume of the feeds the states are getting.

Again the data system is established for analytic purposes, not for paying claims. The data are statistically anonymized, encrypted, and/or deidentified before they are used. Back to the map, the gray areas are ones where there's been low involvement in APCD, but again that could change rapidly depending on healthcare reform at the national level. NADHO is a national organization that closely tracks the details, developments and even governance and funding for all these states. About NADHO, it was established in 1986 to promote the uniformity and public availability of healthcare data for policy research market purposes. Almost all our states have some form of hospital reporting system including emergency department but with the gap that hospital data only provides, we've seen an explosion of all-payer data collection across the states.

So we convened the experts in a learning network and states, and track all things related to APCDs. Again these states on the map have minimal funding. It's often state-based or maybe seed money from a foundation. So we have to rely on state-to-state exchange. We really can't go out for huge vendor contracts and let the vendor do the work. It is heavy lifting on the part of the state. But the state to state transfer and exchange through the APCD Council Learning Network is powerful. It has helped us turn this map darker blue over the years. It really is a community of practice, and we're working on common issues such as ERISA, 42 CFR part 2 standards. We have a common data layout standard and files, definitions and formats. We work together and CMS is wonderful in streamlining and standardizing a Medicare state research extract process for these states.

What are a few use case examples? They're changing all the time but I listed a few here. More comprehensive risk adjustment across payers. If you think about PCP practices they have a mix of payers in their practice in this new world and those calculations vary significantly in demographics, health status, payer mix, socioeconomic status. When you have a whole population to apply risk adjustment you really decrease the practice level variations when you're looking at total expenditures and you increase the predictive power of the analytics, and providers in one state have high acceptance of this type of method. We see that larger sample size is really critical when you're looking down at the network clinic and physician level. When you take a single payer and have smaller numbers of observations the numbers aren't as stable and you really don't get that wide-angle view of the whole system.

Value-based purchasing is a big deal. We were at a NADHO meeting in D.C. last week and had many sessions on that. Again population measures are important. We saw where one state is using their APCD to really look at low-value services, which are running about 30 percent, and that's a low-hanging fruit for any value-based purchasing program. States are using their IPCDs for policy evaluation pre and post to assess the impact on the population and on the cost.

Support and evaluate payment in healthcare reform. How do we target the hot spots so we're not throwing money at places that don't need to be addressed? Setting control versus intervention groups and seeing how they fare. The total cost of care measure is a big deal. We have a standard measure and multiple states are applying this method to look at their own total cost of care measure but also start looking at multi-state comparisons. We are also working on other multi-state reports with standard measures. At least one state is using their data for coordination of benefits as a way to apply administrative savings across plans, so that all plans can benefit from a common source for coordination of benefits.

Again, retrospective and predictive analytics are a big deal. Opioids, case-managed populations, key diagnosis. While individual measures are good in outcome and costs, we're seeing that population

measures with a larger number of patients, you look at the range and distribution of values. You can break it out by payer and see how populations are served in various types of plans. It's really powerful. We saw some output again that states are using their APCDs in that regard.

This is just a concept on why the data and aggregate data are important. The APCDs are designed for analytic purposes, for policy and research and transparency. So this is a cycle that aggregating the data provides a wide angle view of the system and populations, and that facilitates an understanding across stakeholders using a common data set. Each stakeholder is used to looking at their own data, but when you put it all together and bring those stakeholders around the table, you have some interesting conversations about why they're seeing the variation that they're seeing, and it's powerful. I've been involved in many of those conversations. That conversation and understanding supports decisions, both market and policy, and leads to actionable interventions. Which again in the cycle promotes evaluation, and that evaluation usually ends up we need better data, cleaner data and more data. That leads to linkages and the cycle never ends if we're doing our job right.

About APCDs, data quality is a big priority. Jon Huus was at our meeting last week and we had an energetic roundtable and it's gotten a lot of people going here. The APCDs do provide automated quality control protocols, edits, thresholds, benchmarking, and have very clear specifications and reporting requirements. These are developed with input at the beginning from stakeholders including the plans, and they are developed as a community. They're not just presented to the plans and say here's what you need to report to us. It is usually a very detailed and long process to get to the data submission guide.

Testing typically occurs with each carrier, each platform prior to onboarding and going into production and that takes quite a bit of time. So this is not a trivial undertaking. Expensive editing includes front end edits, intermediate edits, and back end analytic edits. The back end edits are ratios, volumes, per member per month calculations by payer and type of service, and also trend edits. We're looking for those spikes and dips that might lead to some missing data or some miscoded data. Payers review and remediate after the initial evaluation and postprocessing edit checks and there are resubmissions required. Again many of these states have laws and reporting requirements and mandates that have compliance built into it and fines and penalties.

Before any analytics can be done, before anything goes public, these known issues of the data and quality control have to occur prior to diving into the data and starting to interpret what's going on. Again, carrier feedback reports are critical to the APCDs and they are provided to each payer for review and remediation, and those carrier reports do go to each payer and each platform. Again I mentioned compliance; that's very important. APCDs usually link the process data back to the raw data files. They have to go back to the source data and verify accuracies. Some of the lessons we learned along the way are payers really don't have the systems we thought they did. They aren't reporting or have trouble reporting outbound 837's or fully populated 837's. So the states are getting extracts or flat files. The payers typically strip down the 837, chop it up, and store it in various warehouses in their system, so they have to extract that data to report to the state APCDs. However, data quality does improve with use and over time.

These are just highlights. Again, there has to be stakeholder involvement in all stages from data collection, analytic methods and reports. All of our states doing public reporting and public APCD development must be transparent and quite open in their methodologies to build the trust so everything is done as a group. Some states are developing data user groups and data submitter groups and working closely with them to improve their data collection and reporting process. It is important early on, and we work with our

states to have clarity, on who gets to use the data, who gets to access the data and under what terms. We have one state on the phone with pretty restrictive data use and access policies and others who have pretty broad access policies. Again that clarity is another way we build trust with the data suppliers so they don't think there is any secret usage of the data they don't know about. Then again, standard and custom reports so that folks can start seeing what's in the data and custom reports for those that have specific data needs are built into the process.

We heard last week, and that's why some of this isn't on the slides, that in many states users really drive the quality improvement in the data—not the agency but the users. The users of the data tell them this is missing, this doesn't look right, this doesn't meet my needs. So at least one state, Maine, has structured a very detailed metadata documentation that provides users information about their APCD data including the defined data validations and edits for each data element so that users can see what is being done and maybe suggest in the future improvements needed.

Another thing important to every state doing this and I'm sure to Medicaid is identity resolution. Not only have the patient—that might be easier for a Medicaid agency than an APCD—but physician identity, it is a big deal, especially as we go to practice and physician-level reporting. I would also encourage Medicaid and APCDs to work with their other government agencies to really think about some shared services such as provider directories. Because we're finding very bad coding coming in on claims for physicians. That really hinders the ability to do what needs to be done as far as value-based purchasing.

When we go out with the slides I will add another link: APCDCouncil.org should be on here as well. This is where a lot of our information is showing up on APCDs including a startup manual for states starting up, some reports, and a use case showcase for how various stakeholders are using APCDs.

As a result of our meeting last week and part of our Learning Network, our members have called for a data quality group. We will ask each agency to designate one participant to come to regular calls. We want to talk about uniform edit standards, standard output reports, so we can start comparing data quality across states with just some frequency distributions and reports. Alternate payment models—different payers defining APM differently. They have different contracts, different contract periods, different payment flows. We're trying to unravel that and have a few states ahead of the game and we're sharing some best practices along those lines. We have states creating specifications and validations to sort of monitor that, and we always disseminate our lessons learned.

Tracy Yee: You mentioned the www.apcdouncil.org website. Is that where the public data is located and can be obtained for those interested?

Denise Love: It's everything about APCDs but not the actual data. Each state in that map governs its own data under its own laws but it will lead you to that state of interest. For instance, if you want to do a study on New Hampshire data, you would have all the information you need to reach out to New Hampshire and find out what they have as a public data set or research data set and how to go about applying for that. There's a lot of information on there, and also how states are using their APCDs. You'll see reports and websites.

Tracy Yee: How do you determine what a low-value service is?

Denise Love: Each community may get together with stakeholders and come up with a methodology. This particular state looked to Choosing Wisely, U.S. Preventative Service Taskforce, CMS quality metrics, and

NICE (Institute for Clinical something in UK). For operational definitions look at NCQA, HEDIS, CMS and the Washington Alliance for codes. They looked and 450 low-value services were identified and 30 were identified as claims-based, so not every low-value service would be claims-based but they looked at the 30 that were claims-based. Then working with their group and other methods, they looked at inclusions, applying Medicare and then APCD exclusions, defining populations to specifications and measures. They ended up with 18 measures. I don't have them but can include that slideshow as this study was done in Virginia. But imaging and preop testing were just examples of some of these that were low-value measures. The population base doesn't...I'm trying to think. Encounter base for denominators. They defined the period they selected. Another one is cancer, over 75 for PSA testing. Then summing up all those encounters and the costs, you can get at what is spent on what the experts have defined as low-value services. This is just one example of how stakeholders come together in a community and decide what their questions are, what their data needs are, vet the measures, and really move that conversation beyond just how many people are getting pap smears but to start looking at where are the savings, where's the low-hanging fruit.

Tracy Yee: One final question to all the presenters. Is anyone aware of what metrics states may use if you're trying to quantify if MCOs are sitting on encounters or withholding them from submission? Any way to look into that?

Denise Love: One things APCDs do is look at other reference databases. They look at inpatient hospitalizations that most states have and just do a rough reference check. They could look at their Medicare data and the APCD and just see if anything looks out of bounds or different. So we won't really know but we can get a sense and then trends over time for those dips and those blips.

Jessie Parker: To quickly summarize takeaways from today's webinar. It's clear to everyone that high-quality encounter data is needed for a multitude of reasons in the state Medicaid environment. State involvement in data collection and use may help support managed care organizations to improve the quality of their encounter data. Other strategies for improving data quality include providing extensive and timely feedback to MCOs, implementing a corrective claims penalty, benchmarking and validating data, developing quality assurance protocols, and editing data.

That wraps it up. Thank you to the presenters for their great presentations. We ask all participants to complete the post-webinar survey. Your feedback on these webinars is extremely helpful. The slides and session recording will be posted on our Medicaid IAP data analytics website.