

Panelists:

Deirdre O'Connor

Deborah Ramelmeier

Please note: The following is a direct transcription and has not been edited.

Deirdre O'Connor: Okay, so I think we are going to go ahead and get started. It's just after the time to start. My name is Deirdre O'Connor and I worked with the National Council on Crime and Delinquency Children's Research Center, which is a national non-profit, it's about 105 years old and the Children's Research Center is a division of that. We've been around for closing in on 30 years and we in that time have primarily focused on supporting agencies, public agencies, and establishing assessments that are evidence-based, research and evidence-based to support improved consistency and validity in decision making throughout the life of CPS case. My co-presenter is Debbie Ramelmeier from Maryland and I will let her introduce herself.

Deborah Ramelmeier: I am the Deputy Executive Director at the Social Services Administration overseeing child welfare and adult services programming for the State of Maryland.

Deirdre O'Connor: And this is session 7.9, Increasing Consistency of Decisions to Accept the Child Protection Report. Just to demonstrate the challenges of inconsistency, we have a different title on the slide and many of you are like, hey is this the right one, is this the right place. So, you are in the right place. We are going to be talking about the development of a screening tool for use here in Maryland, how that was developed, why it was developed, how it was evaluated and then implemented throughout the state and then statewide evaluation.

There just a little bit of housekeeping, towards the back of the room there are handouts on one of the chairs for the data slides that we're going to be looking on at the end of the presentation. There is also available a copy of the article that's in press and you can pick that up if you are interested in reading that. And this presentation is based on that article that will be published in the December, I believe it is the December issue of Public Child Welfare -- Journal of Public Child Welfare.

Okay, I know that this is being recorded. So, if you ask questions, I'll do my best to repeat the question so that if it didn't get recorded, audio recorded then we'll be able to hear it, but feel free to raise your hand and ask questions during the course of the presentation. I think that that may help in the sharing of information, because it's not -- because this is a relatively small group I think that may help in the sharing of information.

We'd like to thank Maryland SSA and Casey Family Programs for supporting this work. Maryland SSA supported it through the desire to implement and evaluate as they

implemented screening and response assessment. Casey Family Programs provided funding for our participation in that and CCD's participation in that, but I don't think that we can underestimate or undervalue the importance of Maryland saying this is worth our agency's time and the most precious resource their workers' time in doing the -- and participating in the evaluation. That's one of the aspects I think of evaluation in child protection agencies that really needs to be raised in terms of importance that the public agencies and the private agencies say it's worth our workers' time to participate in these evaluation activities to gather the -- you know, for a short period of time, gathering additional data or gathering it in a structured format, so that it can be evaluated. So, we are very grateful to Maryland SSA as well as Casey Family Programs for the financing of our participation.

So, we are going to be talking about the impetus for the development of the screening and response tool, the assessment, the actual development of that tool, pilot implementation and then we went through statewide implementation and an evaluation activity for the statewide implementation. We'll also talk a bit about supporting the implementation and answering questions. Debbie also has some data from Maryland from May of 2011, right.

Deborah Ramelmeier: Yes.

Deirdre O'Connor: So, it's recent data that we can talk about in terms of impact. So, with that, I'm going to let Debbie start with the impetus for development.

Deborah Ramelmeier: Okay, everybody can hear me from here. Great. Basically, the reason why we wanted to do this is because that we knew Maryland is a state administered and locally administered system. So, the central office oversees the programming that's done at the local department. All of our local agencies they are state employees, they do report to the governor. They are not locally administered as some states are, but even with that even having policies in place saying this is what Maryland's law requires us to do.

We had vast and I mean vast discrepancies between our local departments on what was accepted in for investigation between jurisdictions and that concerned us, because we felt that especially given Maryland is a very mobile state, people move across jurisdictions fairly frequently. So, we felt it was very important that people in the State of Maryland, regardless of what county they lived in, they should experience our system the same, they should have the expectation that certain behaviors should get the same reaction in every county that if they do a certain act that our departments should either all screen it in or all screen it out.

One of things that we knew though was that where our information for what gets accepted for screening was embedded in our law or embedded in a I think it was a 25-page policy directive that said this is what you should accept. Now, of course, we need to get a phone call how much time do screeners have to pour over a 25-page document in order to remind themselves of well when should I -- what should my response be to this

Session 7.09 – Increasing the Consistency of Decisions To Accept a Child Protection Report

particular thing. So, that is really why the structure decision making tool appealed to us, because it made it very clear and concise for our workers. They didn't have to go into another document. They could get it right on the screen where they are working the referral and inputting the information. It had the information about suggesting to them what their decision should be. So, that is why we decided that we wanted to do that.

Deirdre O'Connor: Okay. One of the other cases that I'll just offer on this slide as well in terms of inconsistencies or the concerns for inconsistencies that we heard as we started the project is that many of the stakeholders in Maryland not just the families, but many other service providers work across county lines. So, the challenge is that they have or that if they were working with a family in Baltimore County and they are more concerned about, and you know, we're concerned about the parents' behavior and they called that county and said hey, this is going on. Baltimore County might respond with an investigation.

Another family with the same behavior in Howard County they would call and Howard County might not respond. So, then the agency is -- then that service provider starts calling state office and saying we should, we should not and it just created an inefficiency in the process because everybody was even at relatively high levels looking at individual situations and trying to mitigate or, you know, facilitate communication as opposed to putting those resources into addressing and serving the needs of the families.

So, from a stakeholder perspective it was also very important that the agency work with the local counties on establishing a baseline for these are the things that we will accept, these are things that we should not investigate. Local counties may have the opportunity to respond to those families with preventive services or with some of the other activities that they were able to develop locally, but they would be thinking that you would not accept would not anticipate it would be investigated for Child Abuse and Neglect.

One of the things to that I think we want to highlight again is that Maryland, the administrators acknowledge that there was a need to improve the screening process and improve the decision making. They were looking for a way to do that. They didn't want to implement a change and just assume that that was going to be a mean better. They acknowledged the fact that change for change sake may not necessarily lead to improvement. So, as part of the plan from the start of how do we make this better was we make a change, we evaluate it, we make a change, we evaluate it, and incorporating that feedback loop was a significant part of the project plan from the very first discussion.

Deborah Ramelmeier: Yeah, and additionally, we did not want to change our law what we do investigate. So, we wanted to ensure that we incorporated what Maryland's current law was.

Deirdre O'Connor: Right. To start that feedback loop, there was a development process, but there was also a pilot implementation and that pilot implementation occurred in counties that were A) most interested in seeing increased consistency among them, because there was significant mobility among their child welfare population and that

Session 7.09 – Increasing the Consistency of Decisions To Accept a Child Protection Report

those counties were also able to participate in the development process. So, we tried to get those that we're going to be early adopters, which is an important piece in terms of any change implementation and child protection and elsewhere that you kind of look for those that are interested in change and start with them, because you are going to be half way through the change process at that point anyway.

And so we did a pilot implementation in evaluation prior to statewide implementation. Then we did some substantial evaluation activities again that we're always part of the project to plan. So, the screening assessment development is based on Maryland law as Debbie said and policy and regulation. It was a 25-page narrative document that attempted to categorize characteristics that should be screened in or screened out, but it was very narrative in structure, so it really was like reading through a 25-page regulation and it wasn't focused. There was no way to pair it down where you should start to look.

CRC reviewed that intensely. We broke it out into a more structured format with some pre-screening characteristics, does it meet three or four basic criteria as a child under 18 or were they under 18 at the time of the allegation, certain things like that that when we reviewed case narrative and actually in screened reports had occurred already, there were times where things got accepted even though it was then clearly identified that the child was over 18 now or there was nobody in the home that was under 18 right now.

So, we did -- we structured the screening decision based on that law. We spent a great deal of time with those counties that we're going to be piloting it and those county -- or the jurisdictions that we're going to be piloting in, the jurisdictions are Baltimore City, Anne Arundel County, which includes Annapolis and then Montgomery County, which includes the Greater DC area. And with those three counties when you -- it was really phenomenal to hear the differences in things like what would constitute neglect, lack of supervision.

Some counties, you know, some of those players said well, if the eight-year-old is home alone for three or four hours after school because the parents working that's not neglect and others were saying, no, we routinely screen those in. So, things like that where clearly different practices we're playing out and making the significant difference both in the volume of investigations that were occurring in the families that were being investigated and then in what occurred at post-investigation.

And in terms like the families that were being investigated, I think that there were two perspectives that we wanted to ensure carry through this process from the family's point of view. One is that having an investigation occurs at varied entries process. Maryland was looking at putting some structure inconsistency in place so that they could then move to the second step of implementing an alternative response process, but you can't get to alternative response until you know that you're consistently looking at the same group of families.

So, having an investigation occur is very intrusive process and we were looking to ensure that that was occurring in those situations in which there was characteristics that met the

criteria that were in Maryland law and not beyond that. And the other piece was that if we could get to just looking at those families there would be a better idea of the resources that were needed to respond to actual allegations and then you would also be able to start saying these are allegations that we may be able to respond to with an alternative response and these are ones that we're going to continue to have to investigate. Sir, you had a question.

Male Speaker 1: Yeah, the three counties that you said you guys had neglected in the State of Maryland the part that how did you go about selecting those counties?

Deirdre O'Connor: Those counties were selected -- do you want to answer that?

Deborah Ramelmeier: Yeah, Baltimore City is our largest jurisdiction. It represents approximately 50% of our case load. So, Baltimore City generally always gets selected. Anne Arundel County had an interest in this. They wanted to be involved in it and they also represented a smaller. We've usually we kind of group our counties as jumbo, extra-large, large, small, extra small, kind of like the crabs and Anne Arundel County is a medium county, Montgomery County is one of our larger counties. They also are dispersed different places in the state. Montgomery County is the DC region, Baltimore City of course the Baltimore region and Anne Arundel County is incorporates our state capital. So, most of it was the desire on those counties from the county administrators that we do want to do this. We're willing to do this and we're willing to give our staff in order to get this done.

Deirdre O'Connor: So, the development process resulted in a developed -- in a structured screening tool with definitions. The screening tool had about five or six items under physical abuse with three or four sentence definitions. It had about eight or nine items under neglect with three or four sentence definitions. It had four or five items under sexual abuse with three or four sentence definitions. I believe medical neglect -- or no emotional injury was the fourth broad category that we looked at that had two items in it with three or four sentence definitions. And it was written when we piloted it in Anne Arundel, Baltimore City and Montgomery. It was a paper document. And then when we did it statewide it went on the web.

Deborah Ramelmeier: Right.

Deirdre O'Connor: So, when we piloted it, the screeners actually had it on -- had a two-page paper form where they completed it and with that two-page paper form there was about 14, 18-page manual that included all of these definitions where they could flip to the definitions. We piloted it in the spring of 2008. We piloted it for about three months. Two months into that pilot, so the workers were trained. They had a short training in about two and a half hour training in the morning and then they would start using the tool immediately after that for all of the calls that they were receiving.

These two of the three counties had dedicated screeners and Arundel had people who rotated into screening and they weren't just dedicated screeners. There were some other

Session 7.09 – Increasing the Consistency of Decisions To Accept a Child Protection Report

folks who are involved. They -- but about -- so they were trained. About two months after training, we had 12 of those from -- we had screeners from all three counties, 12 total participate in an inter-rater reliability test where we gave all 12 a set of actual calls that were documented by screeners, so you would think that it would have all the information that they would need to support a screening decision. We gave 12 workers the same 16 vignettes and asked them to complete the screening tool on those vignettes to see if they would come to the same conclusion given the same information.

We also did a pre and post implementation case file review on the pilot sites. That case file review though didn't occur until a bit latter. But what we saw from the reliability test, inter-rater reliability test from the pilot was that there were several frequently made frequent -- types of reports that are frequently made, like neglect, lack of supervision and risk of harm, physical abuse, risk of harm where harm has not yet occurred or we can't verify that there has been an injury and we're concerned about the reason why, but the behavior indicates that the child is likely to be physically harmed.

So, those were two types of calls that even with the screening tool, the reliability on those types of allegations was lower than what we would want. We took a look at the vignettes that we would expect to get screened in for that or the vignettes that were screened in for that that we did not expect to get screened in for that and found that we could address this low reliability both with improved training on narrative documentation as well as modifying the definitions for these items. So, there were some slight modifications to the definitions of the items, but we also realize that we needed to modify the implementation process.

And so that's the type of that was one of the very simple impacts that really improved the implementation of the screening tool statewide was to say we're going to pilot it and then we're going to evaluate it and we're going to feed that information back into the screening tool itself and see if we need to change it, improve it, or do something else to ensure implementation statewide has a better chance of success with the goal and we know what our goal is. Our goal is to get to consistent decision making on the types of calls.

We are not looking -- the goal was not to get to the same screen in rate across all jurisdictions, because populations are different; its providers call in different things. So, we weren't looking for the same screen in rate across all jurisdictions. We were looking for the individual calls would be screened in or out consistently depending on the information shared.

Female Speaker 1: I have a question.

Deirdre O'Connor: Sure.

Female Speaker 1: What were sort of your ideal reliability rates from these or what did you think of it this is a problem or how did you...?

Deirdre O'Connor: When we did the reliability testing and we'll get into that in just a second, but when we do reliability testing, we primarily look on this type of tool. We primarily look at present agreement, because it's more than one rater; we had a high number of raters; we had a high number of options for the answers. It wasn't a yes/no. So, we were looking at -- so, we've primarily looked at present agreement for each of the items, present agreement for the broad categories did they get physical abuse, did they get neglect, did they get emotional injury, and we also looked at the broad at present agreement for screen in and screen out. We were less focused on that than we were the individual items and the broader categories of abuse or neglect.

Female Speaker 2: Could you repeat the problematic categories again?

Deirdre O'Connor: They -- in Maryland it's broken out into physical abuse, neglect, sexual abuse, and emotional injury.

Deborah Ramelmeier: Or mental injuries.

Deirdre O'Connor: Mental injury.

Deborah Ramelmeier: As we call it.

Female Speaker 2: The ones that you said you had lower agreement on...

Deirdre O'Connor: The item that we had the lowest agreement on in the pilot implementation was risk of harm under physical abuse and I believe it was lack of supervision under neglect. So, we took a look at these three research questions; does the assessment help workers make more consistent screening decisions and the evaluation method for that was to look at inter-rater reliability testing, has the screening decision influenced practice, our workers writing more precise narratives. For that, after statewide implementation we did a qualitative case review and our workers completing the assessments is intended, are they completing it prior to making the decision so that they can inform the decision and we looked at a survey of workers for that.

So, to talk about the inter-rater reliability testing, in this evaluation the numbers here are for the statewide implementation. Once we rolled it out statewide, we went back and did a similar evaluation process with more detail and validity, sort of a broader a bigger evaluation. So, for the statewide inter-rater reliability test, we had 46 workers from 22 jurisdictions. There are 26 in the state.

Deborah Ramelmeier: 24.

Deirdre O'Connor: 24. So, almost every jurisdiction participated in the inter-rater reliability test. The 46 workers reviewed 36 different referral vignettes that were drawn from actual records in CHESSIE, which is the Maryland SACWIS system. Each of the workers completed 12, so we broke the 46 workers out into groups of -- into three

different groups and the each group reviewed 12 vignettes. And again, this is what we're talking about in terms of resources that Maryland dedicated. They -- we did this inter-rater reliability testing via web meeting basically and we created web-based screening tool and we emailed them the document, so it's all -- they didn't have to travel to one location to participate, but we did ask and they did provide dedicated time to do this.

We did the training early in the morning. We asked that they complete all of the vignettes by the end of that work day and so some workers were given half day where they were able to shut their door, turn off their phone, disconnect from email and complete it. Others were told that they should work it into downtime that they had during the day and if they weren't close by 2 O'clock they got that protected time. However, the individual local agencies were able to work with it. They provided protected time for the workers to participate in this evaluation and that's the only way you're really going to know whether it makes a difference for your workers is if you can get the protected time for your workers to participate in the evaluation process.

So, when we did the inter-rater reliability down in bottom you can see that we looked at present agreement, the rate of agreement for the screening decisions and items. We also did look at a Kappa statistic. It was descriptive in nature only. It shouldn't be considered as you know, it's for description because we didn't have the volume that we needed to really use that as a statistic. So, on the initial decision does it meet criteria for one of the screen in items, the average rate of agreement across items was close to 90%, so close to nine out of 10 workers are identifying the same item for screen in and screen out.

There were vignettes where it was lower and there were vignettes where it was higher. There were some vignettes where it was 100%. Everybody agreed on the same screen in decision. There was one where it was very low and that was one that we wanted to go back and take a look at. So, the final decision we also on the screening tool have an override section. Workers can say look it doesn't quite meet criteria on any of these, but we think we need to investigate anyway and here is why. It does meet criteria and we aren't going to investigate anyway and here is why. So that you can see the final decision after override is about, you know, the minimum and average rates are about the same, minimum and maximum are also about the same.

Inter-rater agreement across the individual items ranged from -- the average ranged from about 90% to almost 100%. So, on the individual items, physical abuse, risk of harm, average, you know, the average rates are there and then minimum and maximums are also -- are again from about 50% to about 100%. There were some items where even the minimum rate of agreement was above 90%. There were a couple with a minimum rate of agreement on one or two vignettes was really pretty poor. And even though the average looks fine, looks acceptable on those we again went in and looked at what is it that the narrative said, what is it that we would have expected, is there something that we need to do differently with the tool to get to better rates of agreement or is this an implementation issue. So, we were able to go back in and take a look at that.

Again, we used Kappa for another statistic. This should just be considered descriptive in nature, but you can see that it does fall from the reliability of the 28 items and decisions across intake workers. It's 0.64, which is in the range of good. The 0.76 is just up into the range of excellent. The assessment reliability findings summary of that is that there are high rates of agreement among workers who voluntarily participated in this testing. The Kappa statistic is similar to those of other screening assessments and the findings suggest that it does lead to consistent decision making in the screen in decision. So, when we looked at the inter-rater reliability testing basically what we can say is using this tool, it look likes workers are going to be making consistent decisions. Do you want to talk about the case file review or do you want me to continue doing that Debbie?

Deborah Ramelmeier: No, go ahead.

Deirdre O'Connor: Okay, so then on the case file review for the pre-implementation, we did it for a baseline to get a sense of what narrative look like prior to the use of this tool. We looked at non-pilot agencies because the pilot agencies were already using it. We looked at cases that occurred in September of 2008. We looked at 196 randomly selected cases. We did attempt in this selection of these cases. We did -- we randomly selected them, but we did make sure that they were across all of the jurisdictions. So, I don't know if that's semi-random. I don't know what you would call that. The post-implementation, we did 244 randomly selected cases, both pilot and non-pilot agencies, we looked at cases that were screened in April of 2009.

What we found was that pre-implementation screening supported by narrative versus post-implementation screening supported by narrative, there was a slight improvement, so the screening decisions, screen in and screen out appeared to improve somewhat slightly, but when we look at the allegation types as it being screened in for physical abuse, as it being screened in for neglect, as it being screened in for sexual abuse or some of the other sub-items under there, we saw -- we did see an improvement in the pre-implementation -- from the pre-implementation to post-implementation, in terms of all items, all allegation types supported by narrative. And we also saw that there was an improvement in the response time being supported by the narrative from pre-implementation to post-implementation.

We did, you know again, the purpose of evaluation is not only to say, hey look what a great job we did, but it's also an opportunity to say so what can we do better. Post-implementation, we found that narratives fully documented all the maltreatment allegations indicated in their report and that they matched allegations recorded in CHESSIE. We did see that there was some inconsistent interpretation in some of the neglect allegation definitions and so that was again an opportunity to go back in and strengthen implementation and have a more focused conversation about what is it that's driving some of these inconsistencies, how can we address that as an agency, how can NCCD support the agency in addressing that.

We found that implementation fidelity varied by office. So, some offices were much -- were doing a great job, other offices appeared to be more -- appeared to have greater

challenges in that. There was also a web survey. We had about 70 workers respond to a web survey and what we found was that they found that having the definitions to refer, they were referring to these definitions when completing assessment more than half the time, they did use the override aspect and Debbie will talk a bit about how frequently some use that override. We found that from -- arranged from one to 100 in terms of referencing definitions during the completions, so some workers were using them quite frequently to be able to go back and say does this meet criteria, is this a call that we should screen in.

We also found that most workers knew the screening assessment was intended to inform the decision and reflect agency policy, so they understood the intent of it. They were referencing the definitions again during the completion approximately 60% of the respondents found the assessment helpful, which is something, you know again, to get this information helps the agency know is this really why we have, you know, is it that they didn't understand well, maybe it's not that they didn't understand. It's that they didn't agree. I mean we got a lot of that in the training to the point that you know it's not that they didn't understand what the policy said, it's that they didn't agree with what the policy said, which is a completely different challenge.

And so understanding that, this evaluation really also helped the state office understand what needs to be the focus of quality implementation review, where do they need to put some of their resources to get to even higher levels of consistency. So, I think I'm going to go ahead and let -- I'm going to skip through these slides quickly and we can come back to that, because I want to let Debbie talk a little bit about some of the data that we have available. Now, these slides, these data slides are the ones that we provided copies of for the handouts.

Deborah Ramelmeier: Okay, this is for -- I'll stand out for this one. This is for May of 2011. The information we got of our wonderful what we call MD CHESSIE or SACWIS system and this is actual results of what we screened in and screened out. Maryland has we basically do about 50% screen in and screen out, that's what it was before STM and it continues to be right around that. Of course, this is only one month, May tends to be a big month for us. Schools are getting ready to go out, so we tend to get a few more referrals in the month of May. We also find that the month of May does tend to get a lot of the -- we've had this concern all year round and we just because of schools are getting ready to end we just want to make sure that you all are aware of it, because we're concerned for the summer.

Any of you who work in child welfare know that those are probably going to be screened out. So, we would expect May to be a higher screen out than you know, let's say you know January, where we are getting more of things that are actually occurring right now versus what it might have been over the summer. So, as you could see it's pretty self-explanatory, this is the -- we do have a pretty high override out rate from 11%, which is higher than what we've been told it should be. When I looked at what the reasons were for override out, it was a lot of yeah let me go to the next one before I do that, so this.

Deirdre O'Connor: Yeah.

Deborah Ramelmeier: This is the numbers that will give a sense of how many referrals Maryland takes we do about almost 5,000 a month. I think that's right.

Deirdre O'Connor: Yeah.

Deborah Ramelmeier: I'm having right, yeah.

Deirdre O'Connor: Yeah.

Deborah Ramelmeier: Some months of course are less, May is a pretty big month. So, this is our override out. The ones that are -- the other-- the things that are desi here are what those the tool actually asks you or the categories in our tool. So, one is non CPSK. So non-CPSK of Maryland is that, you know, it might meet the criterion according to the STM, but really what we think this is more appropriate for is that the family really just need services, so, we're going to send out a friendly worker to say, "hi, we'd like to help you," versus the -- as we said the non-friendly worker saying, "hi, I'm here to investigate you."

Maryland does not yet have alternative response. We are hoping to finally get it in the coming legislative session. We figured four times is like the -- you know when you bring something fourth, lesser your four times, they should pass it right. We shouldn't have to do a five. So, that's part of what we I think sometimes seeing with our overrides out. One of the things adult victim and no children in the care in Maryland, we will not investigate if they were victimized as a child and they are now an adult, we do not investigate that. Our issue is does that the perpetrator have access to children, but because the allegation of sex abuse, it does meet the criteria, but it is a valid override out.

Information for two another jurisdiction again, you know, Maryland is a relatively small state and sometimes people don't realize, you know what, if you are on a street, one side of the street is in one county and the other side of the street is in another county and so we do get calls people just don't know what county the person lives in and so we have those. And then insufficient information to locate child or family, that's about you know 4% and those are those calls where I was in the grocery store and I saw this child and I now have no idea who they are or where they live, and no, I didn't go and get the license tag on the car, so those we have to screen out, because there is no way that we can have our director staff is to where to go to investigate that.

Now, the other is a lot of cases and when I looked at what the other was a lot of it was investigation already open, so it's we're getting multiple calls on the same incident and that the other one was -- that the case is already open with one of either in out-of-home or it's already open in in-home family services and it's not a distinct allegation of abuse. It's generally the neglect investigation. Somebody is calling in just say the house is dirty. We're working with the family, because the house is dirty. We're not going to

reinvestigate that. We've already investigated that the house is dirty. So, we don't need to continue to investigate dirty house. So, that's another reason for the override.

The other thing we found though is that some of our staff are not indicating the other. We thought that that was mandatory and that's one of the things we have to look at to see how is it that our staff are able to move forward without filling out what is the reason for other. So, that's something of course that you know now that you know that we're aware of it that we need to look at the central office what is exactly happening with all those others.

Deirdre O'Connor: Can I just make a...

Deborah Ramelmeier: Yeah.

Deirdre O'Connor: So, one of the things that's worth commenting here again is that the structure of the screening tool allows Maryland administrators now to go in and take a look at not only, you know, is it physical abuse, is it neglect, but some other sub-items that are breaking it down underneath those broad categories. They can also then go back and take a look at information like this and they can take a look at information like this by county. So, they can see and I don't think where you have the data slides in here from the counties, the individual counties, but when we look at override rates, some counties are literally have none, zero. You know, they are not overriding anything. Other counties have override rates of like 20%.

Deborah Ramelmeier: Right.

Deirdre O'Connor: So, it's still the issue of consistency in screening decisions has not completely been resolved, because of Maryland's continued use of data they are aware of that and they are better able to target their resources. So, it's really this concept of that, you know, we have these point-in-time evaluations, but I think that again I just want to highlight the benefit of Maryland's commitment to using data, to continue to monitor and improve and move towards their goal of getting to a consistent screening decision.

Deborah Ramelmeier: Yeah. And I will say that prior to STM, if you would ask in the State of Maryland what kinds of physical abuse allegations that you're getting, I go, we screened in however many physical abuses, so many were indicated, so many were substantiated and that's all I could tell you unless I went and read every record. The structure decision making tool allows to break down the physical abuse allegations. So, we actually can look at and can know what kinds of physical abuse allegations are we getting.

Neglect, I mean neglect is a huge nebulous catch all category. We're able to really look to see well, are most of our neglect cases, is it lack of supervision and if we're finding in some jurisdictions and the local jurisdictions are able to look at this data and so that they can look at their data to say okay, in my county we're getting a lot of lack of supervision cases. So, what do we need to do to respond to that. We've a lot of kids where you know

Session 7.09 – Increasing the Consistency of Decisions To Accept a Child Protection Report

the parents are working and they don't have daycare. So, let's deal with that under systematic rather than a case-by-case.

So, it gives us data. We're very, very strong in Maryland about managing with data. We're -- the one great thing about the SACWIS system is that we now have some data that we can look at versus just dealing with anecdotal data, which is what we have before. So, it's allowed us to make decisions up at the central office and the local departments allowing them at the county level to make the decisions that they need to make. Overriding is a smaller number, a lot, the biggest one is domestic violence, where you know the child, it doesn't meet the criteria according to the definition, but there is things that are going on in the family that maybe whether child is in close proximity and so they're overriding those in. And then we also have our little override of risk of harm, where there hasn't been an injury.

Some of these are -- there is a question of whether or not these really meet our criteria for investigation and whether or not these really should be more of these non-CPS friendly visitor cases again, having this information that allows us to address those issues with those local departments that are making these decisions. And then response time, and I threw this in here, because we never before had any way of knowing well, what is our response time, what percentage of our cases needed to our response, what needed 24-hour response and what are the five-day response. In Maryland, all physical abuse and sex abuse cases must be responded within 24 hours of the report being taken, five-day response are neglects and mental injury. The two-hour response are for the emergencies.

They could be neglect cases, you know, small child unattended right now, we're going to send a worker out right away to make sure that small child is not unattended. So, what the STM does allow us though is that we can actually know what the response time is and it also makes it so that the worker has to consider it like before where the worker might not have thought hey, this isn't to our response. Now, they have to think about the fact, oh wait yeah, this is to our response, because I know I have to make that decision here and it's not necessarily some I'm going to pass on to the supervisor and hope that somebody picks, reads in here and then picks up on oh, you know what, this really is something that needs to be responded to right away.

We found that the override rate for the response time, because the system does suggest also response time, the tool. Our override rate for response time is very low. I mean basically we take the recommendation that the tool gives us for response time. So, I thought that was a good sign that we are responding in the timeframe that we said we wanted to.

Deirdre O'Connor: The one other thing that I will say about response time is that when we were going through the development of the tool, both in the pilot and then again as we looked at it for statewide implementation, what we heard from jurisdictions was that they were considering response time as they categorize to the allegation type, because Maryland has this very strong connection between if it's neglect, it's a five-day response; if it's physical abuse, it's a 24-hour response. There were calls that really fell under a

physical abuse allegation type, but the agency was not -- but it was incident between a teen and a parent, and the teen is already back at school and is complaining about his dad hitting him and so do we really need to get out there today, it can be a five-day response. So, we're going to call it neglect instead of physical abuse.

So, they were really looking at, at least anecdotally what we heard, because we couldn't go back and compare data pre-STM implementation and post-STM implementation. What we heard was that there were times where the agency was considering the response time as they looked at the allegation type and thoughtfully identified whether or not they really needed to get out as quickly as a physical abuse or sexual abuse allegation would have required them to get out. What they are now doing instead is accurately quoting or classifying the allegation as what it needs criteria for and then they have the opportunity to say, because the child is able to self-protect and is in a safe place for the next, you know, 48 hours we're going to decrease the response time to, you know, five days or something like that. So, there are -- they have the ability on the override section to modify the response times on when appropriate when certain criteria on that. Yeah.

Male Speaker 2: You may have said this before and I apologize in advance, is the STM tool built into SACWIS?

Deborah Ramelmeier: It is. I was just going to mention that our -- the STM, okay and I don't know how many of you are from states with wonderful SACWIS systems. Our SACWIS system is not very responsive to changes we want to make, because it costs a lot of money to make any changes. So, we implemented this in January of '09. It didn't get into our SACWIS system until July of 2010, so to give you an idea of how long. So, therefore our staff were doing the STM through a web-based that they put then...

Deirdre O'Connor: We did -- we -- CRC was able to build a web-based application.

Male Speaker 2: Right.

Deirdre O'Connor: In Maryland, I think you had -- we had a link...

Deborah Ramelmeier: Yeah.

Deirdre O'Connor: Within SACWIS, so from SACWIS they were able to get on a link, they were able to complete the questions.

Deborah Ramelmeier: And it could get saved into the SACWIS system.

Deirdre O'Connor: And it didn't -- it wasn't completely integrated. It wasn't seamless for workers, but it was as best we could do in a remarkably short period of time like two months, remarkably and expensively.

Deborah Ramelmeier: Yes.

Session 7.09 – Increasing the Consistency of Decisions To Accept a Child Protection Report

Deirdre O'Connor: And collect data and match data and ensure that workers were at least had the opportunity to complete it electronically. One of the issues early on that we found was that because it wasn't required they were allowed to get on to the next step in data recording on SACWIS or case recording in SACWIS. We had to go back and remind everybody to make sure they completed the tool...

Deborah Ramelmeier: Right.

Deirdre O'Connor: Or used it.

Deborah Ramelmeier: And so that certainly we knew when we were doing the STM tool prior to getting into our system that not everybody was doing it, not everybody was looking at it, it wasn't required, it was really basically supervisors you need to make sure it's done, but we knew because we pulled some data in May of -- December of 2009 and very clearly I was like yeah, we're not -- not all departments are even using this. Now, it is required you cannot move forward into accepting a referral without doing. It is embedded in the system as part of documenting a referral. So now, at least I do know that in 100% of the referrals they are completing the tool though you cannot move forward without completing it. So, that's the one you know at least I do have that full faith that they're all doing it. Any other questions about the data. Yeah.

Female Speaker 3: Just on curiosity, are you reviewing any data on the workers, education or anything else...

Deirdre O'Connor: From the survey, we have that information so those workers that did complete the survey, there the majority of them or I guess the average range of experience was seven years in Maryland, the practice is -- and this varies significantly from jurisdiction to jurisdiction. I've worked in a number of different states and with CPS agencies abroad as well. In Maryland, all local jurisdictions really have made a commitment to putting experienced workers at the screening decision point as opposed to new workers. So, they have experienced staff doing the screening.

The average length of time working for the child protection agency was seven years for those that completed the survey, but it went up from there. I mean there were significant -- there are people that participated in the work who had 20 experience that were doing the screening decisions, which is beneficial and challenging all at the same time. Sometimes, those are the ones who really think, no I've got the right decision and everybody needs to come to me. Other times, they are the ones who're able to think through how it's important, the reason for a change and the importance of consistency.

The survey also asked about education and again in Maryland they are hiring primarily master's levels workers for state agents or for the local department. So, the majority had master's a few, less than a quarter had bachelor's. So the majority, you know, 75% or so of the workers that completed the survey have master's.

Deborah Ramelmeier: Yeah, I mean the -- in order to be employed in the child welfare system in Maryland you have to have minimally a bachelor's degree in one of the health sciences and then we also have training that all staff go through and part of the implementation of this was a full-day training on the STM tool and it was really not even on the STM tool. It was on what are the definitions in the State of Maryland. And for workers that are moving into screening, we do have a training that they have to attend part of moving into training -- move into screening to ensure that they understand definitions in the State of Maryland.

Deirdre O'Connor: And just one of the other thing, you know, again the reason that -- one of the reasons that Maryland started this was that they were attempting to get to alternative response in the questions they were getting from legislators and those that would be approving this are you know voting for this legislative change was what about you know what of these physical abuse cases that you are currently screening in would be required for investigation or which would allow for assessment. They couldn't answer that question. That was the level -- that was where the level of data that they had available didn't meet the queries that they were getting.

But then, they also knew that there was going to be inconsistencies in how they responded, what certain jurisdictions would assess and which certain jurisdictions would investigate. Having the structured process of the screening decision for investigations only will allow for when assessment does come into play, when, when it does. It will allow for a structured process to say these are the calls that should be assessed and these are the calls that should be investigated.

I think previous evaluations I've read of like the dual track process or the assessment process is that one of the big challenges jurisdictions face is ensuring consistency and acuity in identifying families for assessment versus investigation. And so to try and put in play a process that will allow for better monitoring of that off the bat and hopefully some preemptive work to ensure consistency and acuity in that decision. I think it puts Maryland in a really good position to support a strong implementation of dual track or alternative response assessment off the bat.

Female Speaker 4: Have you noticed any differences in the number of removals since the assessment implemented?

Deborah Ramelmeier: We actually have not. We're pretty stable as far as the number of removals. Maryland for the last four years has embarked on something we call place matters what we really focused in on our out-of-home population, but really looking at the back door and where kids are going our foster care caseload has dropped by like 25% in the last four years because of that initiative. But we really haven't looked at the front door, because we had again we looked at the data recently and it is pretty much the same number of kids coming in, but they're leaving quicker.

We also -- in addition to doing this, we also have a family center practice initiative where we do family involvement meetings and so and again it's interesting we haven't really

seen a drop off of the number of children coming in, but what we're seeing is the differences in how we're responding once they get in to then allow them to leave quicker.

Female Speaker 4: You said that most of your perceptions of this I mean let's assume clearly the more creative ones are probably know that this isn't where you need to do something is that rather seamless or can it be seamless or is it really because you know workers would have to have suspicions of new and you know how do you sell it to the brokers...

Deborah Ramelmeier: Well, part of that was you know by selecting the counties that we did. We knew that when we first went forth we didn't want to have to spend a lot of time trying to sell it. And I will tell you that it was when we then went out to do the training and we're bringing everybody on board, yeah we had to do some pretty strong sell. We have some counties that were basically we know we're doing it right, why do we need this, because we know that we're doing it right. One of the great things is that this doesn't take hardly any time at all. It is checks. It is very, very quick.

We did have workers complained about having to go into the web-based for that year and a half until we could get into their system. We haven't heard any -- anecdotally anything complaints from local departments about having to do this now that it's built into CHESSIE, because it is so seamless. It's just another we just added two little tabs on the thing, they write out their narrative, they then go and they click okay, this is what these are, this is you know and it's done and I don't even think it takes five minutes for them to complete that part of it. The time what takes a lot is writing the narrative, which they were already required to do.

Deirdre O'Connor: And on that too, I think one of the things that CRC has been able to do the one of the limited ways that we are able to support the integration into CHESSIE was that when we developed the web tool the definitions are available like you see the question as you would on the printed form and right next to it is a question mark that everybody now knows, you click on that and some information is going to be there to help me.

Deborah Ramelmeier: Right.

Deirdre O'Connor: So, you click on that the definition pops up. It's right there. You don't have to find that 25-page manual somewhere on your desk or under your desk or...

Deborah Ramelmeier: Right.

Deirdre O'Connor: Wherever it may have ended up. So it's right there and they were able to, Maryland is able to get their SACWIS developer to include that feature in the tool when it got integrated into CHESSIE. So, all of the definitions are readily available for workers as they're completing the documentation in the automated environment. They can also then, even if they're sure that this is that, you know even if they're

confident that they're selecting the right allegation, they can click on that definition and see what the characteristics are so they can document their narrative more fully.

One of the big pieces and this is about this is like the second or third hotline tool that I have been involved in developing and one of the big things that we see in reviewing and evaluating the impact is that workers think that they're making really clear -- they believe that they're writing really clear comprehensive narrative enough for everybody to understand why they're making the decision they're making and it is not even if it's two or three paragraphs, it's not the information that helps somebody coming behind understand why it fell on one side or another of the watershed. It really is.

And so, having those definitions readily available to workers having them know where they are hearing from the survey two months after they have been -- after this has been implemented that they're using these on a regular basis is one of the ways that can be used and monitored, but also just be aware of so that they can improve their narrative. And I think that's one of the -- and it's also something that's available like if they are thinking well, I think it meets criteria, let me get that definition while the reporter is on the phone. I'm going to go to that definition and ask that last question that I need to ask to make sure I have the information I need to have and I think that's another piece that we would see will -- you know that we could see, that we have seen in other jurisdictions that maybe some jurisdictions in Maryland are using.

Female Speaker 5: You didn't do any measures of inter-rater reliability before you use...

Deirdre O'Connor: We didn't measure inter-rater reliability, no. We simply had the anecdotal information of this is where, you know, we know that we're being inconsistent. Yeah, so we didn't do any in an unstructured environment.

Female Speaker 5: Like the one of the latest...

Deirdre O'Connor: It could have been, yeah. I mean I'll tell you the conversation having done in some of the trainings, to some of the statewide trainings I'll tell you that there were times -- anecdotally, I'll tell you that there were times where in the room we would say does ever -- you know here is the definition for physical abuse, you know, risk of harm. So, here is a narrative that we pulled out of CHESSIE, does this meet criteria or does it not meet this definition. And if we simply put a definition up there and said here is a definition now let's look it a narrative out of CHESSIE, does it meet this definition or does it not, we get 100% consistency, yes it does, no, it doesn't whichever, you know, whichever one it was.

We get consistency and it meets definition; it doesn't meet definition. When we then had to get to so does this get screened in or not, half the room is like yep, we screened it in, half the room is like no we screened it out. We just agreed on what the definition was and so as Debbie said there were times where we were saying we would -- and so, that's really how we ended up structuring the training was let's look at definitions, let's see whether this kind of case meets it or doesn't meet it. And then I know you may not agree

with it, but you have to follow policy. This isn't the Deirdre O'Connor Child Protection Agency. This is the State of Maryland Child Protection Agency. And so, you have to follow this policy.

And it's also an opportunity to say this is where you use the override, because then at least the state office can start tracking some of those cases or some of the types of cases that are really causing challenges. Workers aren't screening things in, because they want to get more work there in screening things out, because they don't have time to do it. They're screening things in because they're really worried about these kids and their screening amount, because they really don't think we need to intrude in their lives. But when it's inconsistent, when some people are screening it in and some people are screening it out, somebody is wrong. You know, so that's where you need to get consistency before you can really start evaluating a level of validity in decision making.

Deborah Ramelmeier: Yeah.

Female Speaker 6: This has been very helpful thank you. Because Maryland also have a structured safety assessment that you're investigating worker who assists...

Deborah Ramelmeier: We do have what we call are SAFE-C and we also have a risk assessment.

Female Speaker 6: And if you looked at C what the screening decisions how it may sort of describe the build sort of the way around that...

Deborah Ramelmeier: We haven't compared like the screening. We actually did have our friends do an evaluation of our SAFE-Cs and our risk assessment and so, one of the things that and we were not surprised by the results which were that you know what, we have vast inconsistencies in our safety assessments and in our risk assessments.

Female Speaker 6: Has to be in Pennsylvania, so.

Deborah Ramelmeier: Yes, and so we are -- it's actually on our wonderful little CFSR pet, we are getting ready to embark on a revamp of our safety assessments and our risk assessments moving towards a actuarial tool that will because...

Female Speaker 6: [indiscernible] [01:00:20]

Deborah Ramelmeier: Our risk assessment is I mean if -- we thought that our screening was went all over the place, our screening was ridiculously consistent compared to what our risk assessments are.

Female Speaker 6: You know I learned that.

Session 7.09 – Increasing the Consistency of Decisions To Accept a Child Protection Report

Deborah Ramelmeier: Yeah, because it's completely left out to the judgment of the worker. There is no structure really in although it supposedly was structured when it was developed, but you know so...

Female Speaker 6: I actually think that if your screen that you're using now is no longer using so much an area that the checklist would have less biased in words, it may actually...

Deborah Ramelmeier: Oh sorry.

Female Speaker 6: All right. Ideally, if your screening assessment is now more structured with more dropdowns and people aren't writing words that can be really, I don't want to say biasing, but can really get the attention suggestive that it may actually you may find that changes. I think because that was one of things that we looked out was that the screeners sometimes in their narratives for writing words that were really sort of suggestive and priming and sort of priming it, you know, sort of...

Deborah Ramelmeier: Oh okay.

Female Speaker 6: Setting the suggestion, so, I think moving to what you've moved to really probably will help you in your safety assessment.

Deborah Ramelmeier: Right.

Deirdre O'Connor: That's my theory or my hypothesis. No theory, not baseline theory.

Deborah Ramelmeier: It should be on the other side of you. Yeah, I mean we've got a lot of work to do on this. Our risk assessments have been in place since I want to -- I think 1997, is when we implemented. I actually was a supervisor out at a local department when they implemented the risk assessment and started training and it was our local department sitting to the risk assessment training. I've never fought and argued more in my life during the training than during this risk assessment training because we couldn't agree and that was what you know and so we haven't know where we are now regarding our risk assessment. So, we kind of see that as like this phase two is now let's you know, that this is where we're headed.

Deirdre O'Connor: Are there any other questions? Again, so there are copies of the article in press. If you want it, pick one up on your way out. On the handout that we provided, it wasn't really intentional, but maybe it's beneficial. You have contact information for both Debbie and myself. We'd be happy to share a copy of the screening tool if you're interested in seeing it, understand that screening assessments are the most idiosyncratic of the assessment tools that CRC has developed. It's really based on the state policy or local policy in legislation. So, it's an opportunity to take a look at the structure of it, but obviously it changes from place to place. And if there is any other, you know, if you like to stick around and talk at all about anything else, we have some time before the keynote begins, but thank you for your attention.

Session 7.09 – Increasing the Consistency of Decisions To Accept a Child Protection Report

Deborah Ramelmeier: Thank you all.

Deirdre O'Connor: We appreciate you for coming.