NewSTEPs 360: June All-Awardee Meeting
Quality Indicators
June 16, 2016

Careema: Specifically we're going to be talking about the timeliness quality indicators, so we will begin with a brief overview of those quality indicators. I'm going to talk a little bit about the new source document that I hope all of you have seen or will be looking at, and then we will have a couple of speakers who will be presenting on their experiences with those particular quality indicator number five, talking about their experiencing of extracting and using that data. Part of this call was in response to some questions that we had on a previous call, that you all wanted to have some more information about the quality indicators. So we hope that today will answer some of your questions. I'm sure you'll have lots, so we will try and keep the questions to the end of the presentation. Keep them in mind, write them down as the folks are going through their presentations, and we can try and answer them then. Just to let you know, we have muted everybody for now, so if you want to present or you have a question, please, you can either unclick the mute button which is at the bottom left hand corner of your screen, or if you're calling in, simply press star six to unmute yourself.

Great, so just a quick overview on the quality indicators. As you know, we are collecting the eight quality indicators in the new data repository. Quality indicators one, two and five focus on timeliness of newborn screening, and today we're going to focus specifically on quality indicator number five, I believe - the experiences of Texas and Colorado. These quality indicators were initially developed way back in 2011, and so over the years we have been working with you, the newborn screening community, to further refine them, define them, and most recently we worked with the quality indicator working group to put a little bit more into the definitions. We've provided a glossary of terms, we took a lot of feedback from those states who have [inaudible 00:02:15] who are entering data in the repository. We took all of that information, and we have come up with a product that we [crosstalk 00:02:24] I'm sorry, did somebody have a question? Okay, I thought I heard something.

So the version two of the quality indicator quality indicator source document is available. I wanted to let you know that the repository is currently being updated to reflect these new source documents. It is going to be an iterative process, so right now we are focusing on those timeliness quality indicators, which are one, two and five, and we expect those to be ready by June 30th in the repository. QIs three, four, six, seven and eight will be subsequently added on, and we hope to complete that in the summer, so will be getting some additional emails from us and announcements from us as to when those are ready. So without further ado, I would like to take this opportunity to invite

Erica, Brendan and Mike [inaudible 00:03:18] to begin their presentation. If you just bear with me, I'm going to share my screen with their presentations. Erica, Brendan and Michael, you can take it away.

Erica: Thanks Careema. We are going to be tag teaming this presentation today, between myself from Colorado, Brendan from Texas, and Michael from PerkinElmer which is the vendor. Can you go to the next slide, Careema? Great. So what we'll do today is we're going to review each of the quality indicators that are being utilized by NewSTEPs 360. These are the ones that Careema alluded to regarding our timeliness, and we'll talk about each of our states' process or barriers in our successes, as we were trying to extract data from our systems for this, and then we also have Michael Johnson joining us from PerkinElmer, and he is our IT extraordinaire, who will really talk to us about what he had to do on his end, and he can complain if he wants about working with us. [inaudible 00:04:25] goosebumps during the process, and we'll talk about our next steps as well as the discussion, and hope that you all have some questions and that you can join in this discussion as well. Next up slide?

So first off I wanted to give you all a comparison of Colorado and Texas. In Colorado our birth rate is about 65,000 births a year. We also do handle screening for Wyoming, which gets us another 7000 infants. We are both two screen states, so we typically see about 140 specimens per year, and we do utilize PerkinElmer's Specimen Gate as our limb system, and also I like to thrown in the fact that we were also super bowl champions, in case you haven't heard recently. So in comparison, Texas has a much higher birth rate than us. We always think of, wow, how do they do it? With 400,000 babies a year. They too are a two screen state, which means they have 800,000 specimens a year that they have to keep data on. They use PerkinElmer specimen gates, and a little bit of a concern we have for them is that they did steal our quarterback from our super bowl champion. Next slide, please. I'm actually going to turn it over to Brendan, because he's talking on the next two slides for us. Brendan?

Brendan: I'm on mute. Oh there we go, I'm not on mute anymore.

Erica: There you go.

Brendan: Thank you. Okay, and just for the record, I'm in Austin, not in Houston, and outside of Houston, nobody cares about the Texans. They're not American made. Anyway, okay. The overall process, so the first thing we want to talk about is the overall process of gathering all the data and what we saw as what it took to do that. The first that we have on here is, define the QI indicator for your state and align with New Step's QI, and I'll refer to this later "define the definitions." The QI indicators, I worked with everybody, Careema and everybody to refine those, and I know an immense amount of work goes into making them as clear as possible. But as hard as we work on it, it's still difficult

to establish how that actually is going to show up in your particular program. That's the first step. One example of that that Erica put on here was, defining what is an unacceptable specimen for transport and what is for collection. The New Step strategy is two buckets, but we have 12 to 15, sometimes more different unsat codes, and we have to establish which codes is which [inaudible 00:07:31] defining the definitions has to happen throughout.

The next step would be to determine which data fields in your system need to be extracted and calculated, and that can take different people to do. You usually need a program expert as well as a data system expert, and then you have to prioritize and identify the resources to actually complete those extractions and calculations, so fit that amongst all the other things that you're doing and working on. So, next slide.

Erica: So we talked about what that overall process looked like for each of our states. In Colorado, we were selected by PerkinElmer to try this out first with data queries - we had actually put in a request to do this over a year ago, and so for us the task was a lot easier than what Brendan is going to share what he had to do in Texas. But what we really needed to do with Colorado was just provide these definitions and these data fields to PerkinElmer, and then our vendor worked on being the one to extract that data and provide it back to us. So yes, we did hog some of the resources of PerkinElmer. Brendan, do you want to talk about your lengthy process?

Brendan: Yeah, sure. We also started on trying to develop some queries for this quite a long time ago, and I'll talk about later one of our very [inaudible 00:09:04] all of the things that we asked PerkinElmer to do. So we always have to prioritize things that come up that are a little bit more important, like figuring out how we can continue to report CS results, things like that. But I'll come back to that later. So basically what we ended up needing to do in order to accomplish this is that, again, like we said, we had to define how we're going to pull the data, and we determined we were going to pull everything by receipt date, because that's what we could do. Then we had some existing data queries that we identified that we could use to more or less pull the raw data that's needed to do all the calculations for reporting. We basically extracted all that to Excel. There's some additional statistics that we had to manually enter into this Excel spreadsheet, and then we coded calculations in Excel to do all the counts for us based on the raw data. So a lot of the calculation that will ultimately done within the lens, we just coded it.

I got to say that it helps me to define our definitions, because if I'm actually in there coding it myself I could go, "Well what about this?" Things like that. So ultimately we set up the template, we extract everything to the Excel document, and it automatically populates the CSV template that we upload into the repository. That was our overall process. So now getting into each individual quality indicator, a reminder of quality indicator one is that this is the percent of

dry blood spot specimens that were unacceptable due to improper collection, and then there's a separate bucket for improper transport. Another aspect, one thing to keep in mind for this, is that this excludes specimens obtained too early for many states, because they either have a [inaudible 00:11:21] protocol or they allow early collection, so again getting into defining the definitions in Texas we don't [inaudible 00:11:29] too early, so that's just one additional specification there.

Erica: So when we look a little bit more at that process or that first quality indicator, for Colorado I and Brendan talked about in our limb system we have 14 reasons why a specimen could be unacceptable, and we had to put those into those two buckets, that collection and transport. And for us, really the only one that perfectly fit in there was our too old for testing. We do have another category when I looked at those 14 reasons with our team that is damaged. However right now, how we collect that data for damaged, we don't keep track of whether that was damaged because of a collection issue, or whether it was damaged because of a transport issue. In saying that, though, we have to make a decision, and we really had to just move on, and when we thought about those damaged specimens, most of the time it appears to be a transport issue. They might come in a ripped envelope, they might come wet because they were sitting outside, or for any other reason that you all see for damaged specimens. But we decided the bulk of those damaged specimens we receive are a result of transport - hence we put it in the transport. But it made us think of, in the future, do we need to better define damaged in our limb system so that we can better query this in the years to come. Brendan?

Brendan: So in Texas, we basically came up with really similar definitions for transport [inaudible 00:13:03] but our approach to it's a little bit different, of course. We assign a code to that transport code of damaged during transport. We'll only do that if we have evidence that comes along with it. So if we [inaudible 00:13:18] a big envelope of specimens and the entire bag is soaked, then we have evidence that yes, it was damaged during transport. But there's other things where we don't necessarily have that evidence, and it may have been a transport issue, but we don't call it that because we don't know for sure and we don't want to report back to the provider erroneously.

So basically we identified those two codes the same, that Erica mentioned, as transport issues, and all the rest basically in reviewing them were in some way shape or form related to a collection issue. In Texas, because of the way I had to pull the data, we couldn't separate out first screens and second screens. Our numerator and our denominator are going to include that. The way we have to go through calculating this is, we have - I'm sure many states have something similar, but we have what we refer to as a global unsat, meaning we did not test it for any [inaudible 00:14:29] meaning that we couldn't complete the entire testing panel, so [inaudible 00:14:36] satisfactory to test for some of the disorders. Basically our magical spreadsheet accounts all the global unsats and

then converts the ... Then we actually have to manually enter the percent of laboratory unsats from our report card calculation, and then it reverse calculates it back to account, and then we add the two together.

One other thing about calculation, anything with any transport issue reported is lumped into that category of transport issues, but the reality is that in Texas we can have up to three unsat codes for any particular specimen, so it could be too old to test but also caked, crowded and layered. But for our calculation purposes we determined to just count that as a transport issue.

Erica: So the issues and barriers that we encountered is that we really in our system how it currently exists, that we don't have specific enough data at times, and when we're just trying to catch up and put in the data from 2012 it's a little too late for us to change those data, since we're doing a retrospective data analysis of this. We also are a two screen state, and I know I sent a few emails to Joshua and Marcie about this. But the percentages that New Step calculates is really based on your total specimens received, so it's a little difficult, because we know from experience that those second screens in our state almost always have a higher rate of unsats in those hospitals, and all of our educational efforts and all of our things we've been focusing on for NewSTEPs 360 are really focused on those hospital births - therefore we're putting numbers in the data that are a taking us to a higher level, and it's a little harder to track our changes and our successes of our educational efforts.

So in thinking about this and trying to figure out the best way to do this, when we were talking to PerkinElmer we asked if they would design a query that we can run all different ways. It could run on just the first screens, and that might be the data we provide to NewSTEPs 360 on a monthly basis. We could also look at the second screens so that we can compare those first and seconds, but also really [inaudible 00:17:05] a cumulative query where we look at all the first and seconds together, and that's the data we enter into our really quality indicators into NewSTEPs. Brendan?

Brendan: I'll back that up as a second screen state, that we totally see a much higher rate for our second screens being unsat. We have a lot of those same issues in terms of being able to measure our education and how to target the education. But some additional issues that we ran into as I mentioned that we had to make the decision on how to handle, that we can actually have three different codes per specimen - we wanted to count that at the specimen level, and how to handle where it could fit into both buckets. Technically, we had some issues with limbs that made things complicated, and I'm sure Michael could back me up on this. But we have very specific specimen categories for our global specimens and our laboratory specimen, because these are laboratory unsat specimens, and this can make the calculations really complicated to do, and as a result of us not just being able to pull this from existing queries, it made it really difficult.

And the two screen state issue, our first and seconds and globals all had different specimen categories, so it can get complicated there. Another thing, our current queries that are available, we weren't able to separate the laboratory unsats into first screens or seconds screens. So ultimately all this didn't end up being a full barrier, because what we're trying to report at this time is, we're basically determined to report on everything, first and seconds. But long term it would be great to be able to separate those out. And one other thing, very few laboratory specimens actually get a transport code, but it is possible for something to be accessioned and ultimately determined to be too old to test, and then we can't report a result. So the way I had to do the counting on that, those are not included. It's a very small number, but that's just one of those decisions that had to be made.

Erica: And so for us, our success is that we were able to successfully work with PerkinElmer to create this new query, and when we looked to compare their new query with what we had been doing on a month to month report card or just from our yearly data, and even some hand counting, it matched. So it told us that the query was working, and we were used to seeing the numbers that it was spitting out. That was a success. Texas?

Brendan: And I would say that ultimately we were able to get to a good representation of our overall unsat[isfactory] rate. We did have some internal tracking that we were already doing, but I think this has improved on it, and it'll allow us to better track some overall trends. We were able to get the data - how's that?

Oh okay, that's me, sorry. Next would be quality indicator two. This is the percent of specimens collected in the appropriate timeframe ... No, sorry, it is not. I'm jumping ahead. It is the percent of specimens missing state defined essential information upon receipt at the lab. There is, as part of the guidelines for this, there are recommendations for what we should be considered to possibly be, what the state defines, so these are different suggestions to consider. But as we all know, every state collects different information and has different requirements for testing, different things that they need to complete follow up, different testing protocols. Those are just some of the overarching issues with quality indicator two and challenges to it.

Erica: So in Colorado with our process we determined which demographics we would be calling essential, and this matched [inaudible 00:21:44] that we determined for NewSTEPs was the list of essentials that we were currently deeming as essential on our report card, so that came to 14 in all. The difficult part for us is that some specimens as you well all know might be missing multiple demographics, but how we were counting those based on I believe the NewSTEPs guidelines was if one essential out of that 14 was missing it was deemed as missing. And we recognized that based on this criteria of having these 14 essentials listed that the data might appear poor to some states because of the high number, but we're really focusing on our goal, which for us is

tracking improvement over time. We'll never get to that 0 percent or those few percent because of how many demographics we have included, but at least we can see what the trend is, and that's our overall goal.

Brendan: So in contrast to those 14, when we sat down in Texas and we reviewed all of our data elements that are listed, and depending on who you ask, everything is pretty much essential. Actually a success that I'm not going to catalog is, we identified two that nobody thought was essential and we got rid of them. But the way we defined essential in terms of this project was to say that we were going to define that as information that was absolutely required in order for us to be able to test the specimen. And ultimately we determined that to be the date of collection, so we absolutely required a valid date of collection in order to be able to test, and from a timeliness perspective, that was the one field that we identified that would keep us from being able to initiate the test, so actually delay the testing and ultimate release of the results.

Of course there are other fields such as PCP information that may delay things on the back end, or other testing fields that a lot of states require in order to be able to do some sort of testing. But in order for us to be able to get to a released result we said that's the one thing that could delay us, so that's how we made that determination. And similar to our lab reporting, our lab unsats, this was a field that we're already including in our report cards, and so was provide a statewide percentage of how often that field was missing. Basically for the purposes of this project we manually enter that into our fancy spreadsheet and it reverse calculates it back to account that we can report to the repository. And I had something else to say about this whole thing, but I don't remember what it was.

Erica: So for us, a barrier in Colorado - this is one of those things where I probably did not give Michael enough information on the front end, our PerkinElmer vendor. We failed to realize that as we've been pushing multiple aspects of quality improvement in Colorado, we have added an essential demographic that's now being entered into our system, and that's mother's phone number. We added that in mid-2015. So when PerkinElmer thought the query was set to go they ran the numbers to provide them back to us. We saw this extraordinarily high rate of missing essential demographics, and then when they ran 2014 and we saw that reach 100 percent I very quickly at that point realized that the reason why our data was not matching was because we had added a demographic field, so half of 2015 was missing and all of 2014 was missing, so we had to go back to the drawing board, and unfortunately ask PerkinElmer to rerun the query to really accommodate for the fact that prior to June 2015 this demographic didn't exist, but we still wanted that data. So it wasn't that we could just give up, but we had to rerun it in a special way so that we could really have enough data.

The other issue that will come up again for us is that two screen state issue, that [inaudible 00:25:42] going to calculate it both on those first and second screens in our state, so we very much recognize that the PCP offices again tend to be our problem children with not completely filling out their card. Brendan?

Brendan: Yeah, and so in Texas just to reiterate defining that essential was a difficult process, and we could define it completely differently going forward. Other issues that we identified when we were looking at it is, one other item that has the potential to delay our processing is if a specimen is received where the electronic information doesn't match the physical form. We were trying to figure out if there was a way to include that in the count, but we don't really have a way of tracking that at this point. And it's not necessarily missing any essential information, it's just that there's a process flow where we don't have everything we need in order to begin the testing process.

Some other things, with the current queries that we have available, we don't have the capability if we were going to look at multiple missing fields to be able to get a quick count of the numbers, so we provide a statewide percentage by individual field, but it doesn't separate those fields. So that's another challenge that we have, being able to gather that data and possibly expand the number of fields we're considering to be essential. Again, I think this is something that I think they came up with solutions for in Colorado. But it's difficult to figure out how do you handle those data elements where you can't just query to see if it's missing. Because it's possible, and in our case, states of collection, if we don't get it, we contact the provider and nag them endlessly until they provide it to us - and then ultimately our database will have that information. So again from a technical perspective how do you count those? We've come up with ways to do that for data collection, but not necessarily other fields.

Erica: So our success is that the query works and we were able to accommodate for that newly added demographic. It also made us really think of all the different changes we've made, so as we were sending Michael from PerkinElmer to work on this, that we would plan early if a demographic was added so that he would be able to run the query in a year long stretch, get quality data, but at the same time accommodate for the fact that maybe an indicator didn't exist, or a data field that was too ... Calculated indicator previously didn't exist, but making sure that we were able to accommodate for that change in our internal process, and hopefully be able to come up with a seamless query for that year.

Brendan: We've ultimately developed, through doing this we've developed ... Whereas we were tracking some of the information on the date of collection missing, I think that by having it provided to the repository in this way we're going to get some more collective information over time that's going to be really helpful for us for tracking trends. Then we had also identified that deficiency in our system for

being able to track other not necessarily data collection issues that can delay testing.

Okay, quality indicator five. Everybody's eyes hurt now. I like that "oh my," Erica, by the way. Okay. Quality indicator five defines each time interval, and this is what we were talking about before. 5a is the birth specimen collection, 5b is the transit time, and then 5c is different aspects of the specimen receipt to reporting out results, so a measure of basically from the time the specimen is received [inaudible 00:30:10] reported out and then 5b separates that out into the larger scale of birth to reporting out results. I think you guys all know the subcategories. There is the additional QI five indicators of reporting now to medical intervention for infants with confirmed clinical diagnosis, and there's birth to confirmation of clinical diagnosis, and birth to determining result with false positive. But I think for the purposes of this presentation we're going to focus on the 5a through 5b.

Erica: So as we start to tease through all those in Colorado, there were certain aspects of five that were relatively easy for us to accommodate for and pick that data. 5a, we had to make some changes of how we were running the query. Out of the report card that we provide to our hospitals monthly, we actually were just lumping together any specimen under 24 hours into one bucket. But we had to really realign that to match NewSTEPs's bucket of that 0 to 12 hours and 12 to 24 hours, and we were able to successfully do that. For 5b, one of the changes we had made over time is, back in the day - again in 2015 - we have always time stamped our cards as they come in. That means that something received at 3:30 pm might not be accessioned today, but at least it's time stamped and so when they finally get around to assessing that card first thing in the morning, they will put that the card was received 3:30pm today.

Previously when we did our queries and we did our timeliness efforts, we actually had run the queries for whatever reason based on when something was accessioned. That meant, and hospitals definitely were upset with us, because that meant even if they [inaudible 00:32:11] to get specimens today but it was too late in the day today, we wouldn't give them credit until it being received till tomorrow. So we were able to fix that in our system last year, but again as we're starting to think about what this process would look like for querying that transit time, we wanted to make sure that we had the most accurate data as possible, and we had to make sure that we accommodated for that change mid-year so that again we could have a seamless data query for that.

For us in Colorado one of our biggest struggles, and I have to say part of it was my struggle, is how to wrap our heads around reporting abnormals. And had a few different ways of how best to pull that data from our limb system. However that's why those were really helpful to talk to the other states, and for us particularly those states that are working with PerkinElmer to see what data field is going to best do that. Then for us for that final report, when that was

accepted, well that was a pretty easy one for us. We have in our reports a quick, easy field to say when that report was accepted, and then of course it also says when that report was printed. So for the tracking of all results was super easy, but for those abnormals it's still one that we're struggling with of how best to pull that from our system. Brendan? On the next ...

Brendan: Yeah, so our process, so first for that collection timeframe, we had to ... If you think about the category, it breaks it out into initial specimens and subsequent specimens. Now, we had to determine how we're going to define what initial specimen is and a subsequent specimen, and there's multiple ways that that could be done, and ultimately what we said that the only way we're going to actually be able to pull data is to say that as we do in our testing laboratory is, we consider an initial specimen to be anything collected at less than seven days. Now, there's going to be a small portion of second screens for the same patient that will fall into that bucket, but we didn't really have a really viable way for us to be able to separate that out. Basically we said an initial specimen equals anything collected at less than seven days, and a subsequent is anything collected at seven days or greater. We also had to ask ourselves how to handle issues where we have a missing time of collection. Our database provides a default time for that so we can do the most conservative calculations on the specimen, which is 11:59 pm.

I'll say that the vast majority of our first screens at least come in with a time of collection as well as a time of birth, so not a huge issue, but just again a decision we had to make, how to process that. So calculations, we're always going to come out with the number of where the timeframe of collection is unknown, it's going to be zero, because we already said how we define that. The same for subsequent specimens collected at less than seven days. That's always going to be zero, because of the way we had to define the definition. The calculations we also had to exclude unsatisfactory specimens where we have no idea how old that is, because we don't know if it's a first screen or a second screen, so we could have a single category for unknown, but we can't categorize unknown initials or unknown follow ups, because we have no idea. This may not be an issue with other states, if you have a direct connection to vital statistics or things like that. I'm not going to say it's impossible to do these calculations, but I think Michael would back me up that the calculations and the formulas that would have to go into coding this would be extremely difficult.

So for transit, this is much more straightforward. Again, we had to define initial versus follow up, and transit for NewSTEPs 360 is just requesting information on first screens, initials - all the different terminology. Again we had to define the missing time of collection, and ... Let's see. I put two different thing on here, sorry. The missing time of birth is 11:59; the missing time of collection is 12:01. Another thing that we had to determine is that we don't actually track the exact received time in our database, so we had to set a default time based on what our specimen receiving area says that the vast majority of the specimens

come in, so we had to make that decision and say since we don't collect it, we can say more often than not this is going to be at least a good indicator of the exact time. Otherwise we wouldn't be able to calculate on time. So again the unknowns would be zero and we'd have to exclude unsats [unsatisfactory specimens] with unknown specimen age. So that's 5a and 5b.

For 5c, for the abnormals, we did make some definitions, but you can see as is the case in Colorado, we were not able to report that information. So we went through and defined some of the definitions and figured out how we were going to pull this data, and we wrote up some specifications, and then Erica hogged all the PerkinElmer resources and we weren't able to get that query done yet. But I know Michael's planning on doing it, real soon. But we're working on it. So some of the definitions that we came up with was that we were going to be looking at first screens only, and this would exclude ... We would have to exclude cystic fibrosis testing from that, based on our testing protocol, because we do IRT/ROT DNA and we don't actually report something that's abnormal strictly on the first screen. This process, keep in mind that this would exclude the first screens that were normal, but the second screen was abnormal. We at the patient level were not able to track that based on the way we set it up.

We also determined that we would count any out of range results for a time critical disorder. I know that that's something that's still under discussion, but from our perspective we wanted to count, if we had a borderline for CAH, we wanted to count that, with the idea that the measure is to measure how quickly we're getting a potential disorder reported. We also said that we had to count these by determination as opposed to the specimen level because for any individual specimen you could get multiple abnormal results. Those abnormal results are released at different times and the provider contact is at different times, so that's another complication to developing the queries for it. And then we went through every one of our possible results, and based on what the disorder was we looked at every result code that was available and determined whether that would fall into the bucket of time critical or time sensitive. Then we also established that we were going to have to use the currently existing data field of case acknowledgment, because up until recently we didn't have a place to track when the first contact to the medical professional was.

Okay, moving on. Like in Colorado we found the final reporting to be a really simple issue. A lot of those same definitions that had to go into the other things, but actually performing those calculations was very straightforward.

Erica: Our barriers and issues, as Brendan said, easy to query when all those results were finalized, but how do we really identify when those abnormals were reported to the PCP? We currently do a lot of putting information into open text fields, and obviously that's not a good field to query in our system. Sometimes those text fields are not completed by our team until hours later, and as Brendan said, and he talked me through this, but there were definitely borderline results

of critical disorders that we hold onto, so we know it's NICU baby and it was a C3, but it's a Sunday morning - we're not going to call those out to our metabolic specialist because of the idea that it's an NICU baby, so there's times we hold onto those borderline results even though they're of critical value. But in talking to Brendan I probably just have to accept the fact that when we put these all in cumulatively we'll see the data for what it is. And how to treat those disorders like CF for second tier testing was also another barrier we had to overcome.

Also for us, as many of you do, we do not run the same testing menu on a weekend, so we definitely ... How do we capture that? That we were able to call out the critical results in saying that, but it doesn't really reflect when we do the overall reporting of how our data's going to look. We also struggle a little bit with the differences of first screens and second screens, and some of the quality indicator five accommodates for this, regarding transit time and collection, but other parts such as resulting, they're still thrown together as one bucket. So again we treat those seconds differently, because on a weekend all the seconds get thrown to the side and our goal of our weekend staff is to just get those first screens run for those critical results. So when you cumulatively throw those into the same bucket at the end and look at your overall time to resulting, that's a problem.

More importantly though, when you do the birth to resulting, we had more than 50 percent of our specimens reported out after 10 days of age, which makes sense as being a second screen state, and so when we were doing these data queries again we wanted to accommodate for how do we divide those up, because we wanted to be able to really look at that data separately. Brendan?

Brendan: Yeah, and so I'm not going to go too in depth into my barriers and issues, because I think I've touched on these already. The first two have to do with how we would define first screen and second screen, and I talked about how we handle missing time of birth and time of collection, the receive date issue, first contact field. But finally, as I always do in every presentation, I need to talk about the competing priorities for IT resources. So especially for the one and two that was an issue, because it's a complicated query to develop. I keep joking about Erica hogging the resources, but internally as I mentioned before we had a lot of things that we're asking PerkinElmer to do, and so some of those queries honestly kept getting pushed to the end of the list due to priority.

Erica: So for us, a success is we were able to really query for most of quality indicator five, and we did do a little bit of nagging of NewSTEPs about maybe improving some of those indicators of five in order to accommodate for us second screeners. I did not get a chance to look at Careema's latest guidance today, but it sounded like maybe particularly for some of those 5d (if I recall specifically) that we really will see the two different buckets for the first screens and the

subsequent screens, so that we can really compare those and treat them differently in our states.

Brendan: And in Texas, I did say for QI five, it really did force us to sit down and say, "Okay let's define these definitions, and what is time critical, and what is time sensitive?" And also recognize some of those facts that there are still deficiencies in our overall system, like it or not, that we can improve upon or need to just document that there could be. We also were able to initiate the whole process of tracking that first contact, which is an important field, and so we were able to improve our database and our program for better tracking of that.

Erica: Now I believe that Michael's going to join us from PerkinElmer to talk a little bit about his experience of working with us. Hey Michael are you on the phone? Because we can't hear you, Michael. We see your lips moving ...

Careema: You can press star six to unmute, Michael, if you're on the phone.

Erica: Do you have a microphone on your computer?

Michael: How's that? Did you get it?

Erica: Now we can hear you.

Michael: All right, am I loud enough?

Erica: Yeah, that's perfect.

Michael: Okay. From a vendor perspective, you heard a lot about Erica and Brendan discussing resources and stuff like that, and how to gather the data, and really that's two sided. It's the same for us as well. We get crunched for that as well, so what we ended up doing was we started out with an initial survey. We took all the QIs and we came out with a survey and we delivered that out to Colorado and Texas, first off. And I think some other states have the exact same survey out, and that helps define what we were talking about, one, with the unsats with the transit and collection times, how to define those specifically. Because each state captures their data differently. Good, bad or ugly, it's what the state wants ultimately, so ... We initially started out with this initial survey, and that at least gets an open dialog going between us and the states, and that leads into the different data points. Everybody captures stuff differently, and that's okay - it just makes it ... I can't just transport it from Colorado to Texas whenever that needs to happen or vice versa. There definitely is some differences between them.

Time allocation. Trying to find time to get into this is tough. As Brendan said, there's other higher priorities on our slate, and that goes for all the states and

also some of the Canadian provinces for us. There are some other things out there. A state wanted to incorporate a new [inaudible 00:48:21] into their timeframe might take priority over a NewSTEPs query, and that's something that we have to balance, and we thank Erica and Brendan and our other states for being patient with us while we work through some of these other things. How to deliver.

This was something that was a challenge for me, because there is a lot of data that's coming out of this, and we have two different systems - screening center, life cycle, and Colorado and Texas; both Erica and Brendan are familiar with life cycle, but we also have a newer product called screening center. The way to display that data, we can write the query on the back end, but to display the data becomes a challenge as well. It's a lot of data points, and Erica, she'll see it when she does get it actually in her office, rather than me just handing her the data. It's quite wide. One line of data with all these in it, it's pushing 100 fields. So that can be a challenge, seeing that, how do you present that to make it good? Because I know they have to plug the numbers in after the fact.

Let's see here. Gathering times for the future. When we started out, we started out down the road of this was going to be a yearly query, and I had a good start with Colorado because I did for them some report cards a while back, and they were the precursor to the NewSTEPs query, so I had a good foundation when we started in with Colorado. But initially it was a yearly thing, so that query is built right now to be run on a yearly basis. Now it's something that we're going to look into that we actually want to ... Do we need to break this down to weekly? I know monthly is on the horizon that we want this to be exportable on a monthly basis. That's something, because that takes some pretty good tweaks into the query itself, to get that data changed.

And lastly the QI five animal - that really was an animal. There's a lot to that one. We've taken it in bits and pieces, start with A, get A squared away. B, once you have your initials and subsequents defined, they spread out a little bit, and it does make it easier. But the time critical and the non-time critical, that is something that we will face across all of our vendors, because some of our ... Not vendors, but some of our customers use a follow up system that we've developed, and we can get maybe some more accurate timeframes to when we contact a physician. Meanwhile we might have to do a reporting time for another, so that's another dive deep into this, and that's something that we're facing right now. Because I know there's some other states that are coming along the pipeline and they're really in here to get, they really need this QI five [inaudible 00:51:29]. So when we talked about time allocation earlier, Brendan hasn't even gotten his final products yet. It's in there, and it's about 80 percent complete, but to get him that full delivery, that's going to take some time to set aside from our standpoint, to finish that up.

There's two of us right now saddled with this task, but it's really falling on my shoulders. We had a good start and then we went to the meeting in DC, and there were some questions there. One thing that I didn't put on this sheet was specimen and newborn based. A lot of the numbers, they get intermixed there, between do we want a specimen count or do we want a newborn count. Again, that's another thing to look at is, how do you gather that upfront data? Because we can go at it by looking at newborns, and then you have to figure out, you put another complexity together - if you're getting it by newborns, then you have to say, which newborns had more than one specimen associated with them? Two screen states, they're obviously going to have two. Whether you count a second one that comes in when the first one was an unsat, and that second one that comes in, is that truly the first now? Those are things that we're still [inaudible 00:52:59] out, and I think even going forward down the [inaudible 00:53:01] I don't know if it can be 100 percent.

I think that goal is maybe more to be 95 percent, because there is always going to be those little nuances that you [inaudible 00:53:11] but for me, that's kind of where we stand.

Erica: So for the next steps for our NewSTEPs data, I think the best thing for us is just to continue to have these discussions of how best to capture the elements, particularly for the animal, the QI five. Talking to Brendan's been super helpful in this process for me, and talking to Michael and seeing what he's doing for Brendan, and I even spoke to some folks from Florida who were struggling with it. They also used PerkinElmer, so we had some commonalities to compare, but because of all the things on PerkinElmer's plate they were waiting in line for when it was their turn to work closely with PerkinElmer to tease all these issues out, so they were trying to do it on the front end, so it was nice to talk to them as they were trying to also talk about what elements they were pulling and how they were going to calculate based on what data they had in their system. We've continued to provide feedback to NewSTEPs when questions and issues arise, and always ... You all have been great in getting back to us right away to say this is what we think, or we talked about it, or this is the issue, and that's been helpful.

For Colorado is, with our next steps is we are just waiting some of the additional adjustments to the NewSTEPs data query, or the NewSTEPs data element so that we can then work with our folks at PerkinElmer and our specimen gate to get us to that completed data fields for QI five particularly. And then as Michael talked about is getting that into a monthly data query that I can run and then provide that data every month to 360. We'll also have to rerun those previous months that we've already entered into our 360, because we'll have improved data quality with these queries being created. Brendan, do you want to talk about your next steps?

Brendan: Yeah. I see our next steps as, we are still working on getting the C/D one and two, so we have a query that's in the works for that and a solution for that. Our next steps are to get that finalized and get that working out. But then I think beyond that what we would like to do is use the experience that we've developed in coming up with these formulas and work more with Michael to look at what he's got for us so we can align it. And ultimately what would be totally awesome is just to be able to put in the dates, click the button, and extract that CS base. That's what we're working towards. I think we have one more slide. We're going to talk about lessons learned. Erica put this on here, but I think this is the most important thing for not getting lost in the weeds. I love the quote, but what if an NICU baby with a borderline time critical result got held over the weekend and it was actually their second screen not their first screen ... There are so many unique scenarios, but what we're trying to capture is our overall trends and how we're doing in the big picture.

Like I do with many of my projects is I tell people, let's not look at the one percent of the time, let's try to get data that gives us a good picture of 95, 99 percent of the time. Other lessons learned, aggregate data with goal ... I just said that, sorry. It's important to keep in mind that you may not be able to necessarily compare apples to apples with other states. Your goal is to see how your state is, how your trends are. We saw in the comparison of some of the essential elements fields - they're not going to be the same. It's okay to leave things blank in NewSTEPs while you're waiting to improve. We don't need to get as much data in there as we can, and it's important to accommodate to changes for changes to our evolving data elements. We're constantly changing those, and everybody's adding things to better track timeliness. Ultimately the goal is to strive for quality.

You really want a good perspective of how you're doing in these things, and a query is just going to be very straightforward, and if it's designed properly it'll just give you hard data that you can work with. Erica, did you have anything you wanted to add to lessons learned?

Erica: No, the only thing I wanted to add was to thank Michael for all of his hard work for PerkinElmer. He was new to PerkinElmer when he got assigned to help Colorado with our revamping of our report card that we were working on for COIN, and so right place at right time for us, so that's why we were able to continue this and talk what he had learned from the report card as one of his first assignments has now become that he's the guru at PerkinElmer for working through these new sub quality indicators. We didn't leave a lot of room for discussion, I apologize Marcie.

Marci: We have 27 seconds, go!

Careema: That was awesome, thank you to all of you. It was a great presentation, lots of food for thought. [crosstalk 00:58:42]

Marci: Yes, just to extend Careema's thank yous, thank you to the presenters today, and to all of the work all of you have put in in [inaudible 00:58:56] these quality indicators, and to PerkinElmer and all of the vendors really in partnering with us to help get these quality indicators out and get the data out in a systematic way. Not only is it helping to get the data out, so eventually Brendan will have that button he can push and data will magically come flying out of his computer, but it also is pushing us to really think critically about what the data elements are, so we are being systematic in collecting them. Right now apples to apples might not be exactly true in some of these cases when some of the state definitions are different, but we're getting very much closer to being able to do that, so thank you all for your diligence in looking at these data.

Careema: So with that we are at the top of the hour. Thank you very much for all of you calling in today. It was great to see many of you, and we encourage those of you who weren't on the webcam to be on webcam for the next time. We'll talk to you soon.

Marcie: Thanks everybody.

Brendan: Thank you.

Careema: Okay