ALLIANCE FOR INNOVATION ON MATERNAL HEALTH

Video Transcript: Practical Use of Statistical Process Control Charts

Inderveer Saini [00:00:02] Thank you Izzy, all right. As you all know, our session follows the same flow, beginning with the introductions, then the guest speaker presentations followed by the Q&A session. The topics for today's educational offering is monitoring and reporting data from QI initiatives. Then we'll be hearing from two of our states who have been signed up for the report outs this session, Oklahoma and Wisconsin. And lastly, before closing out for today, I will be sharing the upcoming community planning session updates with you. Not a nice picture, but my name is Inderveer Saini. I am AIM's data specialist, and I'm joined by Isabel Taylor and Dr. David Laflamme. Again, if you have any questions about these communities of learning, please feel free to reach out to us at aimdatasupport@acog.org. Now a special welcome to our faculty members from the Dartmouth Institute for Health Policy and Clinical Practice. We have Dr. Oliver. Who is an Associated Professor and Daisy Goodman who is Assistant Professor of Obstetrics and Gynecology, a Practice Nurse Midwife and Certified Advanced Practice Addiction Nurse. And the topic again for today's educational offering is monitoring and reporting data from QI Initiatives. Without further ado, we'll hand it over to Dr. Oliver and Goodman.

Dr. Daisy Goodman [00:01:21] So good afternoon. Thank you so much on behalf of Dr. Oliver and myself for inviting us to speak with you today about the practical use of statistical process control charts to monitor and report data from QI initiatives. Next slide, next slide. So Dr. Oliver and I do not have any disclosures. Next slide. And for acknowledgments, we would like to say that some of the material presented here has been collaboratively developed for courses for the Dartmouth Institute for Health Policy and Clinical Practice, and also for the Department of Veteran Affairs National Quality Scholars Program for which Dr. Oliver is core faculty. And I myself was a fellow. And we'd also like to acknowledge the first people and also the land, Wabanaki, which is where current state of New Hampshire is located and from which we're both presenting today. So thank you. Next slide. So our learning objectives today, so after attending this session, we hope that you will be able to interpret a basic statistical process control chart, select the appropriate charts based on data characteristics, create and interpret those charts for both continuous and proportions data sort of basic control chart literacy. And then to apply SPC interpretation to clinical improvement scenarios. And we'll be giving some examples as we go drawn from the bundle metrics themselves. And then, to think about the considerations that we might have for tailoring a visualization to a specific audience and give an example of using those reports to inform and motivate change. And this is our agenda. I will turn it over to Dr. Oliver.

Dr. Brant Oliver [00:03:23] Good afternoon, and I love that picture of me. It was before I had gray hair. It's wonderful. Thanks for including that. And it's a pleasure to be invited to be part of this and to work with Daisy on this presentation. We have an ambitious agenda over the next 50 minutes. We want to cover as much as we can. And we also have videos of all the core content embedded in this slide set for a deeper study and later reference. So let's move right into it. Next slide, please. So in the first part, we're gonna introduce you to the concept of statistical process control or SBC and cover the arguably most commonly

used SBC chart, the XmR or I chart. Next slide, please. So for those of you who have had experience with run charts, you can think of it as a foundational element and a precursor to SBC. So, just as a refresher, a runchart is a time plot or line graph that has a measure of central tendency added. So, in this case, we've got data over time. On the x-axis, there's time points, and on the y-axis there's a measure of interest, and the data is varying up and down over time, and of the red line through the middle is the median, the middle-most value in that distribution. Adding that median turns the time plot into a run chart. And there's what we call a big three elements in assessing a run chart. The first is the median performance level, and the second is the maximum range, the distance between the highest and lowest points in the distribution, and the third is the variation type, this common cause versus special cause variation that we'll get into a little bit more as we talk about SPC. So, run charts remain a very useful and powerful tool for basic improvement measurements. So, even though we're going to talk about SPC today, don't forget about run charts. They're very helpful. Next slide, please. Now the reason why we started with run charts is that SPC charts are quite similar with a few slight tweaks. It's still data over time, although some SPC charts can be used cross sectionally. We'll be talking about data over-time applications today. And you can see that it looks very similar to a run chart, except with these dotted lines on the top and bottom. And these represent control limits. If you think of a regular distribution and sort of the 0.05 p-value level where something becomes significant, you can think of points that go outside these control limits as similar. Those lines are set at three sigma deviations from that central line, that's almost three standard deviations. So the chance probability of a point being outside of those red lines is very low. And so that's what's added to a statistical process control chart. Another thing that's different is that the center line is the mean or the arithmetic average of all the values in the distribution, rather than the median, which is a central point. So this is the anatomy of SPC chart. Points over time, a mean going through the center, and upper and lower control limits. Next slide, please. There's a big three for assessing SPC charts, very similar to the run chart. The first is to assess the mean performance level. What's that centerline value, and is it as an acceptable level for what you're trying to work on or does it need to be improved? The second is the range or precision. Lean Six Sigma calls it precision. It's otherwise known as range. That is again, the biggest distance between that top point and the lowest point in the distribution. How wide is the variation that's occurring? And the third is the type of variation, common cause versus special cause. Next slide, please. So I'm gonna move into the anatomy of the XmR chart. You'll notice here that there are two components to it. The first is called the X chart, or in some textbooks it's called the I chart, or individuals chart. And that refers to the fact that each one of these points is an individual observation, hence the term I chart. So over time, those dark blue points are varying above and below that central line, which is in blue. The central line is the average. And then you have the upper and lower control limits, which are almost three standard deviations above and the center line. And then on the bottom chart, there's something called the moving range. Each one of those points on the button chart is the absolute value distance between successive points on the top chart. And we'll move on to the next slide to see if we can illustrate that for you. So let's start with the green line and circle. That green, that circle point on the lower chart, that's on day two on the moving range chart is the distance between the first two points on the upper chart. And the purple circle point on day three on the low chart is the difference between the second and third points on the chart and so forth. And over time, we get these moving range points going across on the lowest chart. And that's important because the average of those points is used to calculate the control limits on the upper chart. Next slide, please. We won't go into math too much today, but if you're very interested in calculating those control limits, the formulae are here. You basically take the average or the central line and then you add the lower average, or the blue line on the bottom, multiplied times 2.66. And you either add that to get the upper

control limit, or you subtract it to get the lower control limit. So again, for the sake of time, we won't stay too much on the math, but the videos go into this in more detail if you want to calculate the control limits. Next slide, please. This chart here gives you a summary of the types of SPC charts that are used for what we call variable level data. That term is an improvement-specific term. You could also say interval- or ratio-level data or continuous- or continuous light data. These are the different approaches that can be used in improvement SPC. And the one that we're talking about today is the first one, the XmR chart here. And it just gives you a nice summary that you can keep in your back pocket about the characteristics of the eye chart or the XmR chart. We also give you videos that talk about the X Bar S chart, which is a variation on that that can be used in certain circumstances. Next slide, please. Now, before we go into more detail on the analytics, we're gonna take a step back and make sure that we recognize how important context is. So I'm gonna switch over to Daisy for that section.

Dr. Daisy Goodman [00:10:11] So, SPC can also be used to really explore the importance or the impact of context on whether or not we can create meaningful change. So, looking at change over time is what SPC does very well, but we can use the SPC to also show us how certain things that happened in the context, for example, the onset of the COVID pandemic, for examples. Might have actually influenced the outcomes that we see in a specific data set. So, unfortunately, most of the time we focus on interventions and the quality of the interventions and not enough on the context in which they occur, which can have a whole lot of impact on whether they are effective or not. Next slide, please. So you can click through all the animations on this. So for example, we might be able to study an intervention that or some time of improvement in the environment around this particular fish in a fish bowl fairly easily because it's a very controlled environment. But if we try to do something for the fish in picture number three. It might be completely futile. And I'm sure, unfortunately, we've all probably had the experience of working in an environment where we felt like whatever we tried to do to improve things just didn't work. Slide five, I think will resonate for all of us who've been in clinical practice during COVID, that there are times when you feel that the environment is like truly chaotic and it's very difficult even to keep yourself from being swept downstream. On the other hand, there are times when we feel like we're just sort of swimming along in a tranquil water. And then there are time when there are obvious challenges but we also think that we have the opportunity to overcome them. So we thought we'd just do a quick poll to see for you all in the audience, which environments here do you think would be the most conducive to making meaningful change for these fish? So pick one and there isn't really a right answer, but it's an interesting thought exercise.

Dr. Brant Oliver [00:12:26] And if it's possible, can you move the polls over so that they can see? All right, so people can.

Dr. Daisy Goodman [00:12:29] Oh right, so people can see the choices, right, that's a good idea. So let's see, one was. I'm not even sure, two is the fish in the fishbowl, three is the two deceased fish, five is the dam broken with the water rushing through, and then you can see one and four.

Dr. Brant Oliver [00:13:00] I was able to drag the poll aside on my screen, so that may help others if you're trying to see the slide.

Dr. Daisy Goodman [00:13:09] And we'll take the results whenever you have. Great, so four is a favorite, interesting. So four is not a bad choice. I probably would have picked one. And the reason for that is that one of the things that really helps us create meaningful change is a sense of urgency or the sense of something being not right or a challenge that

needs to be overcome. And although it's certainly true that doing something in the fish's natural environment where things are pretty tranquil and you can implement something sort of in peace and quiet would certainly be a good opportunity and probably likely somewhat successful. But in scenario one there's a lot of sense of something needs to be done for this fish and if we just take down that dam the fish could probably swim upstream in a much more effective way and that sometimes is really essential for motivating change. So an interesting thought around context is, what is the actual sense of urgency for the need to make this change? And often we'll see if we look at an SPC analysis that when there is a bad event of some sort, we see a lot of improvement in processes subsequent to that downstream. So next slide, please. So I just wanted to show a quick example, back to the fasting blood glucose example that Grant had put up earlier. So let's say these are glucometer readings for somebody who's been diagnosed with GDMA2. And you can see here that this person who's overall, so average fasting glucose is above 120. So we're not feeling very good about this right now, right? This might be the average, but it's certainly not acceptable to us. Doesn't get a whole lot of benefit from the diabetes educator visit that they have, except maybe for the day after. But when we start insulin at that time, things do seem to be getting quite a bit better until they have a baby shower. And so I'm being able to annotate the events that happen on sort of the course of this data set. Can really help you understand that you don't need to get too excited about what happened on day 13 because it really was an anomaly, you're not going to have a baby shower and eat chocolate cake every day. And probably the insulin intervention is probably going to be effective and we should just let it play out over time and keep watching. Next slide, please. Back to you, Grant.

Dr. Brant Oliver [00:15:47] You've heard us talk a bit about variation. And so one of the key strengths of SPC is the study of variation to inform improvement in near real time, and that's really important. And in order to assess variation, we need to have some categories of different types of variation. And those categories can help us to drive different types of action based off the type of variation we observe. The first type is called common cause variation. Another way of thinking of that is random or chance variation. This is variation in the data that is not caused by a specific external or internal force or exposure or change per se. It's more just chance variation due to the usual variation patterns in the system. An example of this might be, in practical speak, variation in work commute time due to traffic lights, pedestrian traffic, or parking issues. In the special cause category, you can think of that as non-random or significant variation. And that's caused by some circumstances that can be identified or assigned or otherwise known. This could be such as an accident on the roadway heading into work. Road closure, heavy frost or rice in healthcare situations. It could be due to the onset of a pandemic causing your practice pattern to shift from in-person to telehealth. Next slide, please. This slide may be the most important one on the whole slide deck. We're here to talk about data, but the important part is using it to inform improvement. And that's the whole idea of using this type of measurement, SPC. If you see a special cause variation, the advisement to the improvement team is to identify the cause of that non-random variation. And if it's a beneficial effect. To act to maximize, optimize, replicate it, or standardize it. See if we can keep that going. See if can get it better. In a case of common cause variation, this is a random variation pattern. Something is predictable. No non-random effects occurring. It doesn't mean that we just sit back and say, OK, there's nothing we can do about this. And we just look at different factors. If this is actually a randomly varying pattern, we check to see if the average performance level is acceptable. Is it meeting the benchmark or the guideline or the goal? And if not, it gives us a signal to work on redesigning the intervention or the process of the system to get a better result. The second one is. The variation, that precision I was talking about earlier. In some cases, it might be too wide for the clinical situation you're trying to improve. You may

need to get it more narrow, better blood glucose control, for example. If it's all over the place, that variation may be too wide, even if the average value isn't an acceptable level, in that case, you'd work to narrow that average, that range, I'm sorry, or in Lean Six Sigma speak, increase the precision of performance in that process. So this is really a nice pocket reference for the reason why we care about variation in SPC. Identify the type of variation and then use it to inform how you think about your next steps in the improvement process. Next slide please. Another great thing about SPC is that you don't need a statistician to interpret it. Anybody from the front desk receptionist to the volunteer to a clinician to a PhD can understand and analyze an SPC chart if they're aware of the basic rules. Now there are different rules. The ones we're gonna suggest are the ones that IHI or the Institute for Healthcare Improvement use. But there are others, General Electric has their own rules, Juran developed rules, and West Guard developed rules. And those rules are used in different settings, and some are better than others in certain circumstances. But in our case, the IHI rules are recommended. They're often used in healthcare settings. And basically what they're based on is probability calculations of different things happening due to chance. So how likely is it in a SPC chart that an eight-point cluster of consecutive points above or below the center line would be due to chance? Same thing for six or more points going up or down being due to chance or one or more point outside the control limits being due chance. And it turns out that the probabilities of that are much, much less than P.05, which is what we're used to in traditional statistics. The chance probability of a shift approaches 0.02, same thing for a trend, and points outside the control limits are just under 0.05. So that's the basis in statistical speed for these detection rules. But we don't need to necessarily have that statistical understanding to analyze an SBC chart. What you have to know are the rules. A shift is eight or more consecutive points above or below the mean for a SBC chart. Six or more consecutive points all going up or all going down, or one or more points outside the control limits. Now, you'll notice we put the rules for a run chart on the left. They're slightly different, less points for the shift, less point for the run chart. For the trend, and they also have too many or too few runs. And one of the videos we have included in the slide set gives you a review of run charts, but we won't focus on that today. We're going to focus on the right-hand side, the SPC rules. So remember, eight points above or below the mean or the center line for a shift, six or more points going up or down for a trend, or one or more point above the upper control limit or below of the lower control. If you see any of those, that means there's a special cause signal or a non-random variation present. Let's play with one. We have a length of stay data set here for a health care setting. And we've got data over time organized by quarters, three-month intervals. And the time point or the x-axis value is in the left-hand column, and on the right-hand column is the length of state data, the average length of staying that quarter. And so if you hit the next slide. Because this is continuous or continuous-like data, accounting the number of days or averaging the number days over the quarter, that qualifies for our XmR chart, which is used for this type of data. Next slide, please. So here's an XmR, XmR of that data. This is the average length of stay by quarter, that same data you saw on the last slide, shown in an SPC chart, the XmR Chart. So we're going to do a quick analysis of this together. Remember, the big three is the center line, or the mean, the range, and then variation. So the first two I'll do, and the third one I'm going to ask you for help with. The center line or the average overall length of stay is shown by the blue center line. And I can see the value of that is 5.033 days. The average length of state for the overall time period is five days. And so we can think, is that based on what we know from the literature and what we from our clinical experience, is that acceptable or not? The second is the range. We can eyeball that here. The highest point is in quarter two of 2019. It's about 5.5. And the lowest point is arguably 2017 quarter two, which is just below 4.8. The range would be 5.50 minus 4.80. That gives me the distance of the variation. Now the third one, I'm gonna ask for your help with a poll question. We're gonna check for what type of variation is

present. To do that, we have to look for the special cause signals. A shift, which is eight consecutive points above or below. A trend, which has six points going up or going down consecutively. Four points outside those dotted red line control limits. So let's put the poll question up here and see if you can find all the special cost signals. Tell us how many. For the sake of time, maybe we'll give it another 10 seconds or 15 seconds. Survey says, all right, we're all over the place. We've got 19% said there's just one special cost signal. Over half of us said there was at least two and the remainder said three or four. So let's see what the next slide shows. Using annotation, which is by the way acceptable in QI publications, you can actually circle the special cause signals as we've done here. You don't have to use different colors. We've done this for educational purposes, but you can annotate where you see the signals. And so for those of you who said three special cause symbols, you were correct. This one in green is a shift. That's eight or more consecutive points below the center line. Now it could also have been above the sound. Just has to be all eight of those points are above or below. The chance probability of that occurring is extremely low. So that is a non-random variation, a shift. And then each of the red circles is a point outside of the control limits. Each one of those is its own separate special cause signal. So each one of these in and of itself is an indication that the whole process is not in statistical control or not stable. It is not a common cause variation pattern. Another way of saying it is is one of more special cause signals, it is in a special cause variation pattern, there is a non random variation occurring. And what we would do here is we would look at the special cause signal and try to associate again to Daisy's point back to the context, what is happening during that time. That was associated with this non random change was it due to our improvement effort or intervention or implementation that the damn break did. Did we add more fish to the pond? What happened there that caused that observed difference? And then use that information to inform the next steps of improvement. Next slide, please. So that was a quick review of the XmR chart for continuous or continuous-like data. Again, we've got a video that provides almost a half an hour segment just on that chart. So you'll have more depth that you can go into the video. But for the sake of time, we're going to shift now to the other side of the equation. What if you don't have continuous data? What if if you have yes-no data or categorical data? Or a dichotomous data. What type of SPC can you use for that? Well, it turns out that there's a whole bunch of SPCs that you can use for them. We're gonna focus on the most commonly used one called the P chart. You can think of P for proportions, things that have a numerator and a denominator. Those are proportions, and you can think P chart for proportion. Next slide, please. This is a diagram of what we mean of there being a number of options here. This is just given for completeness. In the red box are the different SBC options is even more than that, but we focused on the main ones. The one in green is the one we're going to cover today. And that's where you have in the numerator a binomial possibility, a yes, no, or a zero or one. Did it happen or not? And in the denominator, you have all the chances that it could have possibly happened or the opportunities. So that green circle is where we'll live for the next few minutes. Next slide, please. So with Daisy's help, we took a look at some of the different types of measures in your clinical area that might fit the different types of charts we're talking about today. And many of them, I learned through working with Daisy, are P-chart eligible. Their proportions are percentages. So it could be C-sections per total births, proportion of patients screened for SUD, percent of patients who had a postpartum visit, etc. So in the numerator, you have the total number who had the postpartum visit. In the denominator, you had the total that were eligible for the post partum visit and the resulting statistic is a proportion and you use a P chart to analyze that. Now on the XmR side, you don't have proportion data. You have either continuous data such as time or accounts of deliveries per month. Or number of C-sections performs or a number of staff completing a procedure. So, there isn't that numerator and denominator aspect to it. It's just a continuous or continuous-like measure.

This can be helpful if you're trying to figure out practically about which chart to choose for what thing, and go back to this slide and apply some practical examples from measures you know about to help you decide. Next slide, please. So here's some technical language around the p-chart assumptions. It follows a binomial distribution, which means that the numerator has a binomial or zero or one yes or no characteristic. And so there isn't a, it happened more than once type of capability. It either happened or it didn't. It could have happened three times. It could happen six times, but the binomial nature of the numerator is, did it happen at least once or did it not happen? And just zero or one, it's limited to that. So what that means practically in the final bullet, keep in mind that a P chart can never exceed 100%. The range goes from zero to 100. If you do a P-chart analysis and you get 120 or 130%, check your data, your denominator and numerator might be inverted, or there might be something else going on there that's not exactly correct. A P chart should never exceed a hundred percent. Next slide, please. Here's an example of a P chart. On the left-hand side, we have the time data in the first column, the numerator data in second column, and the denominator data in third column. And this is a SBC analysis, a postpartum follow-up prior to discharge. So thinking about that practically, you've got the dates running in the second column. This is the number of patients who did not have a post-partum visit scheduled. And in the third column is the total number of patients discharged. So the proportion is the proportion of patients that did not get scheduled for a postpartum visit and is being tracked over time. And just like the XmR chart, we've got the points over time. We've got mean in the center line, which is the blue line. And unfortunately, 50% of patients here were not scheduled for the follow-up. That's not too encouraging. We have some work to do here. And we also have the upper and lower control limits shown in the red lines. But you can see how they're kind of funky. They're not straight. They kind of go up and down a little bit. And if we go to the next slide, there's a little call-out that points out why that's the case. Control limits fluctuate based on the denominator size of each point. So if, say, the point here pointed to by the call out in the green circle has a larger control limit size, it kind of bumps up compared to the point before it, that means that the number of observations in the denominator of that point in August of 2009 is smaller than the number points observations in July. And what's great about that is two things. If you're looking at this, you can just by looking at the control limits, you can see if your sample size is staying the same over time, whether it's getting bigger or larger at a certain time, bigger or smaller at certain time points. And you can also have a little bit of adjustment that's going on for sample size throughout. The P chart adjusts for it. And so it's able to be flexible depending on changes in the denominator over time. Next slide, please. We use the same special cause detection rules for all the SPC charts. You may be wondering to yourself, oh, there are different rules for the P chart than for the XmR. Fortunately not, they're the same. You have the shift, the trend and points outside the control limits. I'm going to hand off here to Daisy who has a great application from a published article in your field, so let's talk through this P chart and make some sense out of it using the SPC principles.

Dr. Daisy Goodman [00:33:37] Yeah, so this is work that we did due to the fact that in New Hampshire and Vermont, the leading cause of maternal mortality is opioid overdose, and that's been the case for many years. And also, sort of in more recent years, we've seen that numbers start to accelerate, particularly during the pandemic. So... This was a work we did to ask all prenatal patients whether they would like to have a naloxone rescue kit in their possession. And we asked them all at their first prenatally visit as part of standard screening process. So here you can see the overall rate of screening success in our clinic was only 73%. You can also see. Because we've annotated where the COVID-19 pandemic shot down began. You can also see that the number of prenatal visits, the denominator, right, that we had really shrank right after that second point, blue point here

where it's annotated as pandemic begins. I mean, you can tell that because you can see that the control limits have expanded, meaning that we have a smaller sample size. You can also note that as a result of all of our clinic processes being completely disrupted, which I'm sure you can all relate to, we have two special cause signals there, which refer back to sort of mid-March and April of 2020. So definitely a huge contextual factor that really influenced our effectiveness at getting naloxone out to our patients. But then we recuperated, and we had a fairly steady state around here where we're kind of circling that center average line. And you can also see that we started to see sort of a more normal number of new OP visits as we get towards the right side of this plot. Next slide, please. And then this is our success in actually providing naloxone access to pregnant people who had a diagnosis of opioid use disorder. And here you can see that despite a number of different efforts, we've really kind of stuck at this 50% average rate. So we tried a checklist that was really effective when we first started, but then we sort of settled into getting it out to about half of our patients who have this diagnosis. Definitely room for improvement. I look at this chart and I say, So here we have a common cause variation, right? There are no special cause signals here, but it's really not, our average performance is really not where we would like it to be, even though it's very stable.

Dr. Brant Oliver [00:36:31] You know, Daisy, just a question here. You know we haven't really highlighted how real time oriented these charts can be. We're focusing on showing examples that have the whole data set, but when you were actually doing the work, how were you following this? Were you waiting until the end to look at it or did you do something different?

Dr. Daisy Goodman [00:36:50] Yeah, I know. So we're really fortunate in that we were able to abstract this data. Well, the data on screening rates, we actually get from our analytics Institute and their electronics screeners. So, we can get those on a monthly basis as a report and we do. So I would just enter them into QI macros, which is the data, the SPC hardware software that I use. And sort of on a monthly basis and be able to follow it, which could really help us make decisions as we go along. And then these here on this record are hand abstracted actually, but we're working on an automated process for that.

Dr. Brant Oliver [00:37:37] So we're doing okay. We've got 18 minutes left. We're gonna try and give you a quick introduction to two very important options that you have when doing SPC. One is called fixing control limits and the other is called splitting control limits. You don't always have to use these, but just like other statistical analysis where you can have options to say, stratify your analysis or use. Robust control limits or sandwiched regression approaches, similar in SPC, where you can use different manipulations to increase capabilities to detect certain things. Same thing here with fixing and splitting control limits in SPCs. Next slide, please. We've offered an algorithm here that if you need a sort of something in your pocket guide, for at a glance how and why to consider a fixed or split limits analysis. So this is recommended as a quick reference. And it's a simple flowchart. It starts in the blue zone at the top after you do your standard SBC analysis, like the one that Daisy just showed you, you have some considerations. You can A, just leave it as it is, and in many cases, that's just fine. And in other cases, you can choose either the red or the green option. Now, the red option is used when you want to increase the ability or sensitivity to detect a change compared to a known baseline period. And the right hand side is once you detect a change, it allows you to study what's occurring after that change to see what the new characteristics of performance are after the change and whether that change is maintained or sustained over time. That's an important thing in improvement. You make a change, you see an improvement and then the next question is, can I sustain it? And a split limits analysis is one way to start getting information on that. And there are some criteria that are required for each way you go on the algorithm. If you want to fix the

limits, you have to have a couple of things. You have to a stable baseline, so you need a good, good estimate of what happened before you did your intervention that you can compare to. You have to have a known exposure as well, whether that's something that is identified happening in the environment, whether it's the policy environment, natural environment, etc., or it's something you designed, an improvement intervention or an implementation, etc. Has to be known. And the third element is, yes, we want to really maximize sensitivity to detect that change as soon as possible. On the green side, we have to an either a identified special cause variation signal, such as a shift or a trend, present suggesting there's been a change. Or again, to Daisy's section on context, there are cases that simply understanding the context gives you rationale to say, yeah, it is definitely a new system. Like the COVID pandemic is a heck of a good reason based on context to not wait to detect the shift. You just know because the whole practice fundamentally changed. In my area of multiple sclerosis, it changed in a month. We went from full on-site care to full telehealth care in less than a month, that definitely met criteria for saying we've got two different processes here. And we could make a split limits analysis based on that criteria number two. And then the third is we want to make sure that we need to do it in the first place. Do we want understand the different characteristics of that process after the split? If all three of those aren't met, you probably don't need to do a split-limits analysis. So let's talk about fixed limit analysis in a little bit of detail. Why it's called fixed limits is that what it does is it sets the center line or the average based on the average value of the baseline period. And then it keeps it frozen. And what that does is it can affect the control limits as a result because you've fixed that center line. And what does is compare everything that happens after the baseline to that baseline rate. Now, to do that, it requires a reliable baseline. In an SPC speak, it is a process that is in common cause variation, something we call in statistical control. And you have to know the chronology, you have to know when the baseline period is, when it starts and ends, and that there's enough points in that baseline or enough observations in that baseline to see if that baseline is reliable. And of course, the third bullet, the whole reason for doing it is that it increases sensitivity to detect the effect post intervention. Next slide, please. So here's basic, very basic criteria for a stable baseline assessment. You need at least 12 to 15 observations at a absolute minimum, ideal to have 20 or more. As you have less points, you get a more type two error risk. So if you get over 20, you can definitely be convinced that, okay, yes, I've got acceptable type two error risk, it gets bigger as we go down towards 12. But 12 is still a fairly reasonable type 2 error. And in some practical improvement situations, you can't wait for 20 points on your baseline. You want to get going after 12 or 15. It also has to be in statistical control, meaning that the variation of those points in the baseline period has no shifts, no trends, and no points outside the control limits. It's all in common cause variation. If that's the case, you've got a stable enough baseline to use for fixed limits. Here's an example of a baseline analysis that meets criteria. We have 20 points in this case, so it's the ideal. And we have no shifts, no trends and no points outside the control limits. We have a baseline that's in statistical control or common cause variation. That is a reliable baseline that can be used for fixed limits analysis. Next slide, please. So here's a example of a P chart looking at video visits, the percentage of virtual visits completed by video over weeks, over 30 weeks. And you can see here that using a standard analysis, so this is not using fixed limits. All of the values here are contributing to the center line. We can see a large shift identified circled in red, so nice use of annotation here, and it's labeled, start of the intervention. This is encouraging that an intervention started here, although what's interesting about it is it seems like the improvement started to happen right at the time it was started, so the intervention was extremely fast-acting. Or something else was going on that shifted all that performance upward. So that's using a standard analysis. Now, you know, being a great implementation scientist, Daisy might ask the question, well, geez, what happened if we

fixed the limits here? I wonder if there was something else cooking here. So if we go to the next slide. Here's a fixed limits analysis using the baseline rate. Notice that it's a little lower. The overall average in the previous slide was 0.37. Now it's down to 0.32, which is what the average of our baseline analysis was. So we fixed the control limits at that and it's maintained throughout. And what that does is it increases sensitivity to pick up special costs much sooner. Now it may not look like a dramatic difference between the two, but if you're following this in real time, Daisy's team would have picked up that special cause variation right there where it's red circle. She wouldn't have had to wait the additional time to pick up the shift. So from an improvement perspective, she's able to act on that faster because she's picking it up in week 20 rather than later on in week 30. So that's an example of how the fixed limits approach can really help you in real time. If the goal is to detect things earlier, and like you saw in this example, the baseline and intervention periods were known. So again, this is comparing the two. The top is the standard analysis without fixed limits and the bottom is the fixed limits analysis. Special cost signal picked up day 21, eight days earlier than, although I think it's weak, but in any respect, picked up eight weeks earlier or eight days earlier depending on what scale you're using there. Significant advantage from an improvement perspective. Next slide, please. Now let's move to the other one, split limits. So this splits a single SPC chart into two or more pieces and those two pieces have lives of their own. Stories of their own. They're no longer connected. Each process has its own interpretation of variations characteristics and the key advantage of this is to compare the two to each other. And what is the new characteristics in the new process? I made a change that made things cause a special cause signal and there's an improvement. What's the new mean? How much has the average improved? What's the new range of variation? In the glucose control example, did the range of variation get thinner? Did it get better or did it go in the other direction? And also, is the new process in common cause variation or is it varying non-randomly? Again, is there more special cause variation suggesting that it's not sustaining? It's still unstable. Now this is one area where you could spend a whole day arguing on rationale for splitting control charts. We don't have the time to go into it here. I'll just simply say that there are two camps and both are reasonable. The first is called the empiricist or empirical rationale. This is making a choice to split control and then it's based on an observed special cause signal, especially a sustained one such as a shift or a trend. And that's an empirical observation. So you're not using your knowledge of context, you're waiting for the data to tell you when to split it. Now, the other is a pragmatic rationale which uses context to drive your decision. And in certain cases, there is very strong context evidence suggesting that you should make the split. The important part is to state at the outset what approach you're basing your split on and to acknowledge the risks and benefits of it. The empirical rationale has the strength of being based on statistical probability, but has the weakness sometimes of being slower than it could be if there's a case that is a strong case from a context perspective that you should make the split. The opposite holds for pragmatic. It has the strengths of following the knowledge of context to make a decision to split the analysis. However, it has the risk if you don't know enough about the context or misjudge the context, you could assert that a significant change has happened when in fact statistically it has not. So there are different strengths and weaknesses to these two approaches. Just be aware that they exist and the field still debates them quite vigorously. Both are currently accepted practice. In the videos accompanying this, go into more detail about it. So here we are with a back to the video visits P chart. This is the unsplit chart and we see the sustained special cause signal, right? So it meets the empiricist or empirical criteria right away to consider a split limits analysis. And if I was part of the team, I'd really want to, I'd be interested, it's like, wow, okay, this is actually an increase, a non-random increase in the utilization of video visits. That may be really important to learn more about. So I could consider splitting the limits right at the beginning of that shift. Next slide please.

And what that reveals is something like this. We can see the first process that goes from weeks one through 20 with an average of 32%. And then the remaining process, the second process after the split has a new average of 45. And it doesn't quite yet have 12 or 15 points, but so far what we've observed is no special cause variation in that data after the split. So, so far, the new process is showing a higher average and is sustaining so far. Now, you can't thoroughly conclude that until you reach at least 12 points after the split. That's a minimum required to assess statistical control, but it's encouraging. So if I'm the improvement team following this along, so far, so good that this average has gone up, if my goal is to have more telehealth visits. Next slide, please. So we've got about five minutes left for a quick example on applying this to maternal morbidity. Over to you, Daisy.

Dr. Daisy Goodman [00:50:49] Sure, so this is a familiar, I'm sure to many of us have read this great paper by Dr. Main. And colleagues on reducing racial disparity and severe maternal morbidity from hemorrhage. And you can see just a nice use of splitting the limits here after the intervention of the quality collaborative being implemented. So just an example of how this can be used with a very large data set as well as in sort of in clinic or in hospital analyzes. Next slide, please. So one limitation, of course, to SBC analysis is that one can really only plot one data series at a time with it. And so when we think about the need that we have to stratify or disaggregate our data by race, ethnicity and payer to at least to be able to really detect where there is variation in how. Process improvement is applied and or how it is effective to improve outcomes for people. It's important to think about how we might be able to use this analysis and look at differences between groups. And one way, which may seem somewhat clumsy, but is visually, I would say compelling is to look, is to stack your, to plot different data sets for different groups. Sequentially, but then to stack them and align them. So you can really see here that our whole sample had an average rate of whatever it was at 3%. And you can tell that it's a larger sample because of the narrowness of the control limits there. Marking where the intervention was, it appears that for the entire sample, it looks like the intervention was probably somewhat helpful, although I'm not really sure if we have any special cause signals there, maybe towards the right end. But for group one, you can see that a larger sample size rate, that there were some differences in the way the intervention impacted, and the overall average rate there had dropped to 2%. Whereas if you look at group two, you can say that there is a higher average rate and much less precision. As well, and of course the control limits are larger because it's a smaller smaller group size. So this is one way that we can look at disaggregated data using SPC analysis. It's an area that I'm really interested in. And I think needs a whole lot more exploration from a methodologic perspective by people much smarter than Next slide, please. So in summary, SPC is a powerful tool for analyzing the success of maternal health interventions. We can utilize it to track implementation success. So how we have implemented meaningful change, as well as to track our outcomes. Annotation can be helpful, and I would say is actually crucial to understand where the barriers and facilitators of that change are, where they fall, and then to interpret them. As potentially things we would like to replicate or things we need to overcome. Variables and approaches can be tailored to specific audiences, of course, and then visualizing change or lack of change over time can be an important motivator for implementation. So I really encourage teams to use this type of analysis because it does allow us to really in real time see how we're doing with our improvement work. So I think we are at time. I don't know if we have time for questions, but I'll open up the floor. Thank you so much for this opportunity to talk about a topic that Dr. Oliver and I both really love.