Cyberseminar Transcript

Date: October 21, 2019

Series: Spotlight on Mental Health Centers of Excellence

Session: Validation and Acceptability of PTSD Primary Care Screen for DSM-5

Presenter: Michelle Bovin, PhD; Paula Schnurr, PhD

*This is an unedited transcript of this session. As such, it may contain omissions or errors due to sound quality or misinterpretation. For clarification or verification of any points in the transcript, please refer to the audio version posted at* [http://www.hsrd.research.va.gov/cyberseminars/catalog-archive.cfm](file:///C%3A%5CUsers%5CVHASLCMyersK%5CAppData%5CLocal%5CMicrosoft%5CWindows%5CTemporary%20Internet%20Files%5CContent.IE5%5C73RXHSNF%5Cl)

Dr. Paula Schnurr: Good afternoon everyone. This is Paula Schnurr. And Michelle Bovin and I are excited to present to you today the results of our HSR&D funded study validating the PTSD Primary Care Screen for DSM-5. We’re going to divvy up the workload as follows. I’ll provide the introductory remarks and then Michelle will take over talking about the study and then I’ll come back on at the end for discussion and we’re aiming to end by 10 of the hour so that we have time for questions. If we could have the next slide. We wanted to start out by asking about who the audience is, that will help us in tailoring our remarks today.

So there’s a poll and I will let Rob take over on administering the poll.

Rob: Thank you. That poll is up and the question is, what is your primary role in VA? Student, trainee, or fellow, clinician, researcher, administrator, manager or policy-maker, and other. And by the way Michelle it looked like you may have had a little bit of trouble forwarding the slide. If so, just click into it in full-screen mode and it will tell windows that you’re interested in that slide being able to forward. So we have 84% of your viewing audience already making their choices and that usually levels off somewhere between 70 and 80 so I think that’s great and I’m going to go ahead and close the poll and share the results out. And let you know that 16% say that they’re a student, trainee, or fellow, 22% clinician, 38% largest number researchers, 9% administrator, manager, or policy-maker, and 16% other. And I’m going to go ahead and close the poll so we’re now back on the slides.

Dr. Paula Schnurr: Thank you very much. Well it seems like, I hope we have something for everyone because we have a variety, people who have a variety of roles in the VA. So on this next slide there’s a bit of background about PTSD and the DSM-5. So the Diagnostic and Statistical Manual for Mental Disorders was revised in 2013. The PTSD diagnosis, along with many others underwent substantial changes. Next slide.

So the changes included the classification of PTSD into a new category called trauma- and stressor-related disorders. Previously it had been thought of as an anxiety disorder. But PTSD, adjustment disorder, and other disorders caused by events are now in their own category. criterion A which is of the stressor exposure criterion that was changed to be more specific around a sudden death of a loved one such that the death had to be due to a traumatic cause. There were also clarifications around occupational exposures. In addition the avoidance and numbing cluster was broken into two separate criteria. We’ve known forever pretty much that those symptoms did not hang together that well and now they are in their own separate clusters. There were wording changes to 10 of the original 17 symptoms and then 3 new symptoms were added to make the total 20 including symptoms that reflect other changes in cognition and mood. And Michelle if we have the next slide.

That will show the audience all of the criteria that are required. So stressor in criterion B intrusion was the same, avoidance criterion C had been part of the C cluster previously, criterion D is a new cluster that included the numbing symptoms of the prior criterion C in DSM-4 but also additions for things like guilt and shame. Arousal and reactivity is criterion A [sic] that used to be called hyperarousal and the duration remain the same so that a diagnosis of PTSD is not used until the symptoms persist for at least 30 days. The symptoms have to cause clinically significant distress or impairment and not be attributable to the effects of substance use or medical conditions. These are pretty standard requirements for all diagnoses. So if I could have the next slide.

Also to try to give you some context for why it’s important that we validated this screen. PTSD affects a large minority of Veterans and especially those who use VHA. So according to the Northeast Program Evaluation Center, in a 2018 report 11.6% of Veteran users that year carried the diagnosis of PTSD. The prevalence is higher among our female users than among our male users. So 17.7% of the women and 11% of the men carry the diagnosis. The prevalence is also higher in the youngest Veterans, the OEF/OIF cohort, 26.8% of them have PTSD. There the numbers shift to being a little bit higher for the men, a little bit lower for the women but the prevalence is much closer in the OEF/OIF cohort. Now besides the prevalence in our population PTSD is important because it’s associated with impaired functioning, lower quality of life, poor physical health, comorbid psychiatric disorders, and especially depression and substance use. So it’s a highly prevalent condition that has wide-ranging effects on our patients. Can I have the next slide?

So screening for PTSD in VHA is crucial. The goal of screening is to identify those at risk for PTSD or with undiagnosed PTSD for the purposes of intervention. We don’t screen to find cases, we screen to find cases to move them into appropriate care. VA has mandated PTSD screening now for some time. In the first five years after separation and then every five years thereafter with the vast majority of the screening being conducted in primary care. Next slide please.

So we have been using, up until a few years ago, the PC-PTSD which was the DSM-4 version of the scale that we’re talking about today. It reflects the work of Annabel Prins at our Dissemination and Training Division in the National Center for PTSD, also with Rachel Kimerling who’s at that Division, and Paige we met who’s formerly of VA, it was a collaborative effort. And the scale has served VA well but with the new criteria it was important to determine whether we needed to revise this scale and then the optimal way of doing that. So if I could have the next slide.

First I wanted to show you what the, what you’re looking at here is the PC-PTSD-4 what we had been using. And it begins with an opening question. In your life have you ever had any experience that was so frightening, horrible, or upsetting that in the past month you, and then each of the four symptom clusters is asked about just as a yes/no. So cluster one nightmares and thought about it when you didn’t want to, that’s reexperiencing. Tried hard not to think about it or went out of your way to avoid situations that reminded you of it, avoidance. Constantly on guard, watchful, startled that’s hyperarousal. And then numb or detached from others, activities, surroundings that was numbing. And so two and four had reflected the old criterion C. Now if you, can I, Michelle if you could back up for a minute. What I wanted to point out because it informed what we did with the new scale is that this frontend question is kind of complicated. You have to hold in your head have you ever had any experience that was so frightening, horrible, or upsetting so you’ve got to bring up that idea and then combine that with questions about how you were reacting to it in the past month. Hold, if you can just hold that thought, I will come back to saying how we adjusted in a couple of slides. So Michelle, next slide.

So we revised the scale to reflect the DSM criteria. We now call this the PC-PTSD-5 and we did that by adding a trauma prompt and also a cluster to, a question to represent that new criterion about changes in cognition and mood. And I’ll show you what that, the new scale looks like on the next slide.

So first of all we started with a trauma exposure question. We had heard complaints from the field, a lot of them. We had data and we actually had some user experience testing indicating that people could get confused about the kind of events we were asking about. And often in practice the people might be react, answering to non-traumatic events such as a job loss, divorce, and so on. And so what we did in revising the scale, this had nothing to do with DSM-5 but it had to do with enhancing the performance, is including a prompt question in which we defined the kind of events we’re talking about. So sometimes things happen to people that are unusually or especially frightening, horrible, or traumatic. We start with that frame and then we give a list of the kind of events we’re talking about, accident or fire, physical or sexual assault, war, et cetera to try to scope it to these severely traumatic events and not just distressing stressors. And then we ask have you ever experienced this kind of event. So we’re trying to specify and simplify here yes or no. If yes, then you continue with the questions. Now the first four questions are the same as in the PC-PTSD but the fifth question includes guilt and blame and so that and self-blame because we felt of all the new items that were added to the DSM criteria this would be especially prevalent and would capture those kinds of symptoms. So if I could have the next slide to tell you what we learned about it going into the study we’re talking about today.

So we did several pilot studies. Anabel Prins led in effort and published a paper a couple of years ago on our pilot work. Showing that the new scale had strong psychometric properties. But we felt it, we did not have a gold-standard criterion for validating it. And so that’s what this new study was about. We wanted to validate in a primary care sample using a gold-standard instrument with the goal of identifying the optimal cutoff. On the DSM-4 version it had been a three out of four. With our pilot data in which we used for example one of the gold-standard criteria we used or one of the validation criteria we used was the Mini-Mental Status Exam which is not the same as say a SCID or a Clinician-Administered PTSD Scale. Those, we had some evidence suggesting that three might be the optimal cut point out of the five and that four might be the optimal cut point out of the five. And so based on that we conducted the following study and Michelle is going to tell you about our findings.

Dr. Michelle Bovin: Thanks Paula. So, yes so to talk about our study in particular. We had three goals for it. As Paula mentioned we really wanted to identify the optimal cutoff score for this updated measure, specifically in primary care using a gold-standard PTSD interview. This setting is important because as we’ll talk about in a little bit, the population that you’re interested in, where you screening, and the purposes of your screening really determine what cut score is going to be the best. In addition we had two other aims. First of all we wanted to look at whether there were optimal cutoff scores that would differ across different subgroups of interest. We looked at a number of these. For the purposes of this talk we’re just going to talk about gender differences that we had examined. And finally we wanted to gather some initial information about how Veterans thought about the new instrument, were the instructions clear, were the questions clear, and ways in which they would prefer to have it administered.

So we were able to recruit 495 Veterans across two VA sites, one in Boston and one in Palo Alto. We brought them in within seven days of a primary care appointment. We wanted to have a consecutive primary care sample which we were able to accomplish. However some of them couldn’t meet with us on the day of their appointment, so we scheduled them to come back after that to complete a number of self-report measures. We were able to get 429 of them back to participate in a phone interview. The requirement for the phone interview which was session two was we really tried to get them in within 30 days of session one because we did want to be able to compare some of the measures that we had given at session one with those from session two and make sure it was in roughly the same timeframe. Although for our main aims and the data that we’re going to present today, all of it was based on, or 99% was based on data that we collected from session two. And we were pretty good about that goal regardless. Most of them came back in about 12 days with only really 3 participants outside of that window.

Different measures across the different sessions. At the first session we asked about demographic characteristics. We asked about a number of other screening measures. These were measures that we administered. So we weren’t pulling it from their charts but their identical to what is administered in VHA, so the PHQ-9 which is the Patient Health Questionnaire 9 or the depression screen, the AUDIT which is an alcohol screen, MST screen we used the identical questions that asked in VHA. We also gave the PCL-5. Then at session two the participants called, we called them back and first a research assistant administered the PC-PTSD-5 orally. After that he or she administered a number of questions about the screener. The research assistant then transferred the call over to a PhD level clinician who was blinded to their results on the PC-PTSD-5. That assessor then administered the MINI Suicide Module to ensure safety, the Life Events Checklist for DSM-5 which allowed them to pick the worst event, and finally the Clinician-Administered PTSD Scale for DSM-5. This is the gold-standard measure for assessing PTSD based on DSM-5 criteria.

Our final sample, so these are the folks who both came to session two and had complete data from session two, were 399. You can see mostly male, mostly white, and about 7% identified as Spanish/Hispanic/Latinx Ethnicity.

We were interested in determining how the folks who were in our complete sample of 399 differed from those 96 folks who either didn’t come back or who dropped out right at the beginning of session two. We found that on the vast majority of measures that we looked at there were no differences. In fact there were really only three that popped. Non-completers were significantly more likely to identify as Hispanic, to indicate that they were separated from their spouse, and to be employed for wages. And this is pretty consistent with the literature and also with timing that would be required for folks to come back to participate in the second session.

So for data analyses, for the majority of what I’m going to talk about today we used diagnostic utility analyses which is also referred to as signal detection analyses. I’m going to go into that in a lot more detail. And then for our third aim we just looked at descriptive statistics to get a feel for, just understanding the acceptability of this scale to our Veterans.

So because I’m going to talk a lot about diagnostic utility analyses and understanding our results as predicated on an understanding of this type of work I wanted to just back up and give a little bit of an explanation of how we get the findings that I’m going to present. So when you’re conducting diagnostic utility analyses your first step is to calculate these 2x2 contingency tables or crosstabs. You could see at the top I have PC-PTSD-5 which is the test. So basically for each level of the test, in this case Paula had shown you the screen with these yes/no’s and each of the items yes is equal to one, no is equal to zero, you add them up and you have a sum score. So for each level if your score, let’s say was one and anyone who scored one or greater would be a yes on the PC-PTSD-5 and anyone who scored less than a one, in this case zero, would be a no. And that would change as you looked at each level. Similarly for the CAPS-5 which was our criterion or you know, what we used to sort of say this is the diagnosis for the purposes of this study, depending on what rule of the CAPS you use you get a dichotomous variable of those who meet for PTSD and those who don’t. The yes and the no. And this allows us to group people into these four different categories. The true positives who are people who both meet criteria for PTSD according to the CAPS and screen positive on the PC-PTSD-5. The false positives who don’t meet criteria on the CAPS but do screen positive on the test. The true negatives who don’t screen positive on the test and also don’t have the diagnosis according to the CAPS. And finally the false negatives. These are the folks who are not screening positive on our test but who do have PTSD according to the CAPS. And then for each of these 2x2 contingency tables you can calculate measures of test performance and test quality.

So the test performance measures that you’re probably most familiar with are these. The sensitivity, specificity, and efficiency. So sensitivity is the ability of the test to correctly identify all patients with the disease. It’s a proportion of the people who actually have the disease who screened positive over all of the people who have the disease regardless of whether they screened positive. Specificity is really the reverse of this. The ability of the test to correctly identify all the patients without the disease. Or in other words the proportion of the true negatives, the people who really don’t have the disease, over all of the people who screened negative. And then finally the, or who were negative on the CAPS. And then finally efficiency which is the proportion of correctly classified participants among all the participants. So all of the correctly classified folks. The true positives and the true negatives over everybody. And then these measures can be used to come up with various other test performance measures that provide various additional information about the cut score.

Two that folks may also be familiar with are the positive predictive value and the negative predictive value. These again are proportions. So positive predictive value being the probability of having the disease in a participant with a positive test result. And of course you want a high probability that people who have the disease are identified with the folks who test positive. And then the opposite being the negative predictive value. The probability of not having a disease in a participant with a negative test result. It’s worth noting that predictive value is strongly affected by disease prevalence. So positive predictive value increases and negative predictive value decreases as prevalence of the disease in the population increases. Again another reason why you really want to be testing this within your population of interest. Because as prevalence varies so does the potential cut score that would be recommended.

Final test performance measures that we’re going to talk about today are the likelihood ratios. The positive likelihood ratio is the probability of a person with the disease testing positive, divided by the probability of a person who does not have the disease testing positive. And in contrast a negative likelihood ratio is the probability of a person who has the disease testing negative, divided by the probability of a person who does not have the disease testing negative. Now again ideally you would want PP, your positive likelihood ratio the folks who have the disease have a much higher chance of testing positive than does somebody who doesn’t have the disease. So one of the rules of thumb that we use is we like to see positive likelihood ratios which are five or greater which tells us that the probability of the person with the disease testing positive is 30% or more likely than the probability of a person who does not have the disease testing positive. Similarly with negative likelihood ratios we really want to see this to be much less likely for the person who has the disease to test negative than somebody who doesn’t. And so our rule of thumb here is point two or less which ensures that it’s about 30% less likely or more that the person who has the disease is going to test negative than the person who doesn’t.

So it’s worth noting that the problem with each of these is with all of these test performance measures is that they’re uncalibrated for chance. And that means that the agreement between the test which is, in this case the PC-PTSD-5 and the diagnosis which we’re calculated by the CAPS can be spurious. And therefore this can be kind of ambiguous in terms of indicators of diagnostic utility. So in 1992 Helena Kraemer wrote an amazing book where she defined these new metrics which do correct for chance. And these are our kappas that can be used with fixed endpoints. Zero indicates only chance agreement whereas one indicates perfect agreement, not taking chance into consideration. And you can see here I’ve written out the quality of sensitivity, quality of specificity, and quality of efficiency these relate to our sensitivity corrected by chance, specificity corrected by chance, and efficiency corrected by chance. Quality of efficiency is also known as Cohen’s kappa and some of you may have heard of it in that way as well.

So Kraemer thought about these different optimal cut scores as being useful in different ways and having different strengths and weaknesses. Optimally sensitive cutoffs increased true positives and decreased false negatives which is great if you’re trying to cast a wide net and you’re not as concerned with getting the wrong people and you want to be really confident you got the right people. The problem is, it also gets a lot of the wrong people these false positives people, people who are screening positive but don’t actually have the disease. On the other hand optimally specific cutoffs increase true negatives and decrease false positives but they increase false negatives. So these are really good for differential diagnoses not as good for screening. Finally we have these optimally efficient cutoffs. These balance our false positives and false negatives. So they maximize the proportion of correctly classified participants. In other words more true positives and true negatives, fewer false positives and false negatives. Which can be useful if you want to make sure that your false positives and false negatives are balanced but in some cases you might be more forgiving of one or the other.

So for the purposes of our study we examined the five levels of the PC-PTSD-5 against scores of one through five, across four CAPS-5 diagnostic definitions. Lenient which would be just meeting the DSM-5 algorithm that Paula discussed earlier. In other words one reexperiencing symptom, one avoidance symptom, two of those negative alterations in mood and cognition, and two hyperarousal symptoms. The problem with that is that on the CAPS you can hit that with a pretty low-severity score. So other definitions that include severity scores have been tested. We looked at a bunch of different ones, the best fit for the data, which was the algorithm plus the requirement of a severity score of 22 or greater, moderate same except that raised to 23, and finally severe 26 or greater. I’m going to present to you for sort of visually the moderate, but we have all of the others if you’re interested. And we were really interested here in the comparison of two cutoff scores. The optimally sensitive cutoff score with an acceptable specificity. We put that in because otherwise the optimal sensitivity would always be the lowest number. Versus the optimally efficient. And we were interested in determining are they the same, are they different, and if they’re different is one better than the other for our purposes. And I’ll say that in terms of the four CAPS definitions we did find a range of PTSD prevalence which ranged from about 14% to about 20% of the sample.

So looking at this, this is a diagnostic utility table. I’m going to take you through this starting at the left. You can see that about 16.5% met criteria for PTSD according to this moderate definition. That left-hand column are the different levels of the PC-PTSD-5 and would represent the different cut scores. The next column is the percent that has PTSD according to these different cut scores. If we said that everyone who screened positive had PTSD. Then it gives you the false negatives. The percent of those with CAPS PTSD who are not being identified by each of the CAPS scores. The false positives which are the percent of those without CAPS PTSD that are being captured. And then sensitivity, specificity, and efficiency. And you can see here and this is pretty consistent across all of the definitions that the cut score of three is looking like our optimally sensitive cut score. And I’ll get to sort of the details on this in a second. But you can see that this is the highest sensitivity where we have our most efficient is the cut score of four.

Looking at some of those test performance measures we can again see that we’re sort of seeing the same pattern. The likelihood, positive likelihood ratio’s a bit lower than we’d like it to be for a cutoff of three with four being, looking really nice. Whereas we’re seeing the reverse for the negative likelihood ration. And I also do want to point out that you can see that as we increase the level of the test from three to four our false negatives increase and our false positives decrease.

Finally as I mentioned, looking at these quality metrics we can see that three has the highest quality of sensitivity with the acceptable specificity, whereas a cut score of four has the optimal efficiency. And you can also see again how those false negatives and false positives are impacted by shifting.

So across our CAPS-5 definitions, a cut score of three was optimally sensitive whereas four was optimally efficient. As I had mentioned before the cutoff of three reduced those false negatives but increased the false positives. We don’t have our positive likelihood ratio really where we would want it to be but our negative likelihood ratio remains below that point two. And we see the opposite with the cutoff of four were increasing those false negatives but reducing the false positives, better positive likelihood ratio but negative likelihood ratio is a little bit higher than we’d like to see.

So based on this we were a bit torn in terms of which would be best. We can see arguments for both. So we did a bunch of post-hoc analyses to try and figure out if we could get any additional information. First question that we asked was whether adding a PCL-5 severity requirement to a cutoff of three would improve validity. We calculated the best fit for the data was an optimal score of 32 or greater on the PCL-5 and so we looked at three different cutoffs. The cutoffs of three and four as we just talked about and then this cutoff of four and a cutoff of three with folks who also have the requirement of a PCL-5 score of 32 or greater.

And you’ll notice that these numbers are a little bit different because our sample size decreased a little. This is because the PCL-5 was administered at session one and so we have slightly fewer people here. But still again lots of power. And we’re seeing pretty much across the board that that middle cut score that we added in, it really falls in between the two. It’s not as optimally sensitive as a three, not as optimally efficient as a four, kind of hanging in the middle.

And we see that also in terms of our other test performance metrics although it does get a likelihood ratio, positive likelihood ratio that’s ending up sort of more in the range that we want.

But again you know we’re sort of finding across the board it’s falling in the middle.

So again in general the cutoff of three remained optimally sensitive. That was true across three of the four CAPS definitions for the strict definition we actually found that this was the optimally sensitive cut score but it was the difference of a CAP of point seven versus a CAP of point eight, I’m sorry point eight-seven versus point eight-eight. And then a cutoff of four remained optimally efficient across all.

Okay. Post-hoc analyses.

And Paula was mentioning that there may have been a typo there, the likelihood ratio that I had spoken about over here I believe. Right the positive likelihood ratio for four is actually greater than five not less than five.

Second set of, so thank you Paula, and then the second set of post-hoc analyses that we looked at will, was whether the false negative Veterans would be captured by other VHA screens. So we examined false negatives on these other measures that we had given at session one. For the Patient Health Questionnaire-2 the cutoff score used by VA is three so we looked at folks who scored three or greater on that. And for the AUDIT folks both in literature and also in VHA use a cut score of eight or greater. So that’s what we looked at here.

So to take you through this on the top of the table we have the different definitions. So lenient, the data-driven, the moderate, and the strict. The next line are the cutoffs. Cutoffs of three and cutoffs of four for each of these definitions. And then the percentage of false negatives that screened positive on each of these measures. So the percent that screened positive on the PHQ-2. So using lenient as an example, 41.7% of the false negatives using a cut score of three screened positive on the PHQ-2. The next line we’re looking at the AUDIT, 33% of them screened, of the false negatives screened positives on the AUDIT when we used a cut score of three. And then our final line is that the false negatives screened positive on at least one of those two screeners. So we see here that regardless of the CAPS-5 definition or the level of the cutoff that’s used the majority of the false negatives screened positive on at least one of these screening measures that is used by VHA, with the caveat that these aren’t the actual screens but we are certainly using the same metrics.

And so next we looked at, as I mentioned, subgroups of interest in particular for this presentation we looked at men versus women. Men, we have similar to the overall sample prevalence is about 12-17% depending upon the definition. And across all of the definitions we see the same thing that we saw on the full sample. Right, the three is optimally sensitive the four is optimally efficient.

Where we start to see differences are across women. You notice that we have a much smaller number of women but the prevalence for PTSD among women is notably higher. We’re seeing between 25 and over 31% for women. Looking at the different tables we saw that a cutoff of three was optimally sensitive for three of the CAPS definitions. But our positive likelihood ratio dropped way down, it was in the threes instead of the fours. So for our negative likelihood ratio still for the most part would be where we want it to be. However we see that for the strict definition the four was optimally sensitive, we have a good positive likelihood ratio although our negative likelihood ratio was a little bit high. But across all four CAPS definitions a cutoff of five was optimally efficient. In other words they had to have endorsed every single item. And we see huge likelihood ratio, positive likelihood ratios, but our negative likelihood ratios are way higher than we would want to see. So looking at this we would start to think that maybe a cut score of four would be best for women rather than a cut score of three.

And specifically we see that when we start to look at the impact of false negative, on false negatives and false positives across men and women. We can see that when we use a cut score of three we have approximately the same rates of false negatives for men versus women. However when we start using a cut score of four we add about eight percentage points to men whereas we triple in some cases the impact that it has on women. Particularly if you look at the data-driven, the moderate, and even higher with the strict. So we are losing, even with the lenient, if we think about the lenient, this means that we are losing 40% of women who have PTSD because they’re not screening positive with a cut score of four. With false positives we of course see the reverse. The cut score of three we have a much higher percentage of false positives for both men and women and this drops more for men than for women. However although our percentage is higher for women because we have such a smaller number of actual women that this reflects this isn’t actually going to end up flooding the system with a lot more people.

Finally aim three, I want to go through this quickly so that Paula can tie us up and we can have some time for questions. Overall we found that the instrument was well accepted by Veterans. It was rated as easy or very easy to understand by almost the entire sample. A large majority also found answering the questions was either easier or very easy. And this was on a scale from very easy to very difficult. We asked them about how clear the instructions were on a scale of very clear to very unclear and you can see almost everybody said it was very clear or clear. And again the majority said they would be comfortable or very comfortable completing this instrument at a primary care appointment.

Interestingly there were some differences across how Veterans wanted the instrument administered. Folks said that they would be most comfortable being asked by their primary care provider. And they were significantly more comfortable being asked by their primary care provider than they would be completing the measure on their own. Which is quite interesting. And then they were significantly less comfortable being asked by a nurse or another primary care doc that wasn’t their own doctor. They were significantly less comfortable, you can see I put in the T-test here for being comfortable completing their measure, the measure on their own and even more so in terms of being asked by their PCP, so that was far and away the most preferred.

So I’m going to turn this over to Paula to finish up and then I think we will have some time for questions. Paula?

Dr. Paula Schnurr: Thank you Michelle. Using the mute button does have its disadvantages. So what we’ve shared with you today are our initial analyses and essentially what we have found is that the cutoff of three is optimally sensitive and the cutoff of four is optimally efficient. For a cutoff of three we find more true positives, we decrease false negatives, but we also increase false positives, people who don’t have PTSD who are screening positive. We were able to mitigate against this somewhat by adding a PCL-5 severity requirement and so, so you know one, you saw a slide in which we showed you that all of the measures improve if we count it as a positive of either someone who was a four or who was a three and also had a score of 33 on the PCL-5. We also showed you that many false negatives, people who we would otherwise be missing, actually get identified as cases on other VA screens. And we did this analysis to help us further understand the tradeoffs because determining a cut point is all about tradeoffs, how important is the sensitivity versus the specificity in terms of the cost to the system in terms of the cost to the participant, both being overdiagnosed and underdiagnosed. So the cutoff of four would be, from a systems perspective optimally efficient in maximizing the correctly classified participants but it does increase false negatives. May I have the next slide.

Now the data on women were interesting to us and were very important because we had never been able to see a sample that had enough women to perform this kind of analysis with the other studies that have been done on the PC-PTSD. And so it’s possible we would have found something analogous with the PC-PTSD we just didn’t have the numbers. For women we found three would be optimally sensitive cutoff, five is optimally efficient, and, but that when a cutoff of four was used we had a higher percentage of women as compared to men being classified as false negatives. This is suggesting that a lower cutoff score may be useful among female Veterans. I think what I don’t remember or maybe we haven’t done yet is looking at what we can do if we use a three plus a PCL of 33 to perhaps balance out the sensitivity and specificity to optimize the overall efficiency. Next slide please.

So we found that, however that the PCL was well tolerated by our participants and that some of the last data Michelle presented I thought were particularly interesting because what we found is that participants would prefer to have this scale administered by their provider. In my experience this is not the way we have been screening in VA. Often there’s someone like a nurse or some, a care manager, someone else is screening participants. I think in some places this may be done on a written form but in my experience at least PCP’s rarely are doing this screening and at least we know in terms of Veterans’ preference it seems like they want to have a conversation with their primary care provider about this. So I’m going to stop so that we have time for a few questions. And I think Rob will be reading the questions to us. So thank you very much and thank you Michelle.

Rob: Thank you Paula. We do have one or two questions queued up. But let me let the audience know that if you have any question at all that you’d like to ask of Michelle and Paula there is a section in that GoToWebinar control panel entitled questions and you can pull it out and make it bigger if you want. And I have one here and I’ll just launch right in. Were the particular questions out of the five rather than just the sum of the endorsement used for analysis. For example maybe some of the questions should have higher weightings than others and that would help with the optimizing efficiency, sensitivity.

Dr. Michelle Bovin: That’s a really interesting question. Paula do you want to take this or do you want me to?

Dr. Paula Schnurr: Why don’t you start and then I’ll chime in, if I need to.

Dr. Michelle Bovin: So that’s a really interesting question. We haven’t done that. And it’s certainly possible that we could. But the reason that we really wanted to, the reason that it was designed this way and the reason that we sort of capped it, is because as Paula was mentioning for the most part each of the symptoms really represents or each of the questions really tries to represent one of the clusters and so we’re sort of thinking you know the higher chance of meeting the diagnosis is endorsing any of them. But it’s certainly an interesting way that we could potentially look at it to try and figure out if there were differences. First of all across who endorses what and whether there was anything systematic that we could identify about those participants.

Dr. Paula Schnurr: I think it’s an interesting question that we could further explore. We hadn’t considered it previously. I think that if we were going to use other than equal weighting I would especially want to get multiple samples to make sure that kind of finding was robust. But it’s possible you will see a paper on this topic in [unintelligible 48:06] at some point. So thank you for the question.

Rob: That was the only question that we had queued up. Again, if you have a question for Paula or Michelle there’s a section in the GoToWebinar control panel or dashboard entitled questions. You can just type your questions in there and I can ask.

Dr. Paula Schnurr: You know what, I think I want to talk. People can stop me from talking by asking questions. So if you have a question I will stop right now.

Rob: Okay we have one.

Dr. Paula Schnurr: Okay.

Rob: Would the difference between male and female Veterans and rates of PTSD be due to age or what era the Veterans serve? And the person goes on but I think you get the gist of the question.

Dr. Paula Schnurr: So I think we were trying to disentangle that. I mean also men and women have different exposures. And we’re still working through those data. It’s certainly possible that, that gender is also carrying with it many other things that might be a more important determining factor. These data are somewhat hot off the press. I think one thing that is going to be important for us to look at is the nature of whether people have military sexual trauma and so it’s possible that what, because many women do especially in the older cohorts it’s possible that what we’re finding for women may generalize to men who had military sexual trauma but we don’t know that yet. I wanted and this is a great question because actually what I was going to come back to is this particular issue. We are still working through understanding what seems to be happening with the women and how we would make recommendations. Because ideally the recommendations would be simple. And we’ve already presented some complications and not just three versus four but what if you had a three and you added a PCL-5, if we need to the VA will use or certainly could consider using different thresholds for men versus women. But we’re hoping we can be sure of that, that we understand that finding and what its implications would be downstream for the population and I think we’re very actively involved in data analysis now so that we have the best answer on that. Michelle, did you want to add anything to that question?

Dr. Michelle Bovin: No I think that’s exactly right. Our first pass was looking at obviously the, you know our first and foremost was to make sure that we sort of understood what was going on with the overall sample and looking at that across different definitions of the CAPS and sort of conducting all of those post-hoc tests. We’ve also conducted differential analyses across a number of different subgroups. As I mentioned men and women is the only one that we’re presenting today but we looked at this across age and ethnicity and we’ve started to look at it across race, although our numbers are a bit small, to try and understand what’s going on. I think the next step is going to be looking at multiple memberships but we just haven’t cracked that egg yet.

Rob: Those were all the questions that we have at this time. I think Paula, you had some more comments that you wanted to make and I’ll let you know if another [inaudible 51:53] comes in.

Dr. Paula Schnurr: Okay, sure. I wanted to talk about the women but then I wanted to get back to the issue of what we’ve learned from the preference data to the best of knowledge this is the first that we have seen data on our Veterans’ preferences for how screening is conducted. And this is a, although this study was done at only two sites it’s a very large sample for a screening study. So I feel like we have to pay attention to these data and the data suggests a consideration, a reconsideration of how the PC-PTSD is administered and very specifically our data are suggesting that Veterans’ preference would be having it part of the care interaction rather than part of the screening, the paperwork, the assessment by assistants that leads up to meeting with the primary care provider. Now I don’t know if any PCP’s are on the phone. I know PCP’s are sprinting usually from the beginning to the end of a visit. So adding one more thing could be challenging from a workload perspective and we are not recommending that people go out and change practice. We’re simply saying our data suggests that this is what Veterans would prefer and so I think it’s useful to consider that information. Again, in the context of the entire system and all of the moving parts that are taking place in a patient visit.

Dr. Michelle Bovin: Paula I just wanted to add to that. I totally agree and one of the assuasive [sic] analyses that we didn’t present today is that we found that comfort with completing the instrument decreased as severity of PTSD increased. So these considerations may be particularly important for symptomatic folks.

Rob: We had another question come in. What was the reason that the guilt/blame question was not part of the earlier PC-PTSD-4?

Dr. Michelle Bovin: Uh, it wasn’t part of the DSM-4 definition. That was actually representative of two of the three new items that were added to DSM-5.

Dr. Paula Schnurr: So it’s interesting. Guilt used to be part of the diagnosis in DSM-3 and then it got moved out for a while and then it made a comeback in DSM-5. So even though we, guilt is a common symptom in many individuals who have PTSD it had never, it had not been part of the diagnostic criteria for a very long time.

Dr. Michelle Bovin: The survivor guilt was an associated feature but in a screener like this where the, you know the goal is speed, you know being able to gather as much as you can in a short period of time, there just wouldn’t have been room to add it or rationale for it at that point.

Dr. Paula Schnurr: And could I just say I think it also goes along with the recharacterization of PTSD outside of the anxiety disorders framework. Because there was an acknowledgment that there were cases that we certainly thought had PTSD that didn’t seem to have a classic fear-based presentation or even classic fear-based traumatic experiences and that, these more cognitive and mood symptoms were very important in understanding their total presentation. And so I think more than anything we have been learning from the data how important symptoms like this were in this population. And that’s why guilt and self-blame and other symptoms like that made a reentry into the DSM-5.

Rob: The questionnaire replies, fascinating thanks. Thank you for the background and the really insightful and important work you’re doing on this.

Dr. Michelle Bovin: Thank you!

Rob: That’s all we have. At this time. I’m sorry, go ahead.

Dr. Paula Schnurr: I was, it sounded like we were closing and I was just going to point out that we are only an email away and if people have further questions about this, to contact Michelle or myself at the addresses listed on the screen. [ END OF AUDIO ]