

Program Performance Data

Webinar held on November 10, 2020

Watch the full webinar [here](#).

View the future webinar schedule [here](#).

View previous webinars [here](#).

SH – Sandra Hodgson

PS – Dr. Pooja Shivraj

WL– Dr. Wilma Larsen

SH: Hello everybody. My name is Sandra Hodgson. In the interest of time and to respect everyone's time, it's now one o'clock, so we're going to go ahead and get started. Again, my name is Sandra Hodgson. I'm Director of Communications for ABOG. I'd like to welcome everyone to this webinar. We truly appreciate you signing up for this event and joining us today.

This webinar topic today is on program performance data. As a reminder, as it says up on the screen there, the chat functionality is disabled for this event and everyone's microphones should already be muted.

Our panelists today are Dr. Pooja Shivraj, who is our Manager of Research and Psychometrics, and Dr. Wilma Larsen, who is our Associate Executive Director of Examinations for ABOG. I'm going to go through basically our agenda for today. We're going to start out the event by having Dr. Shivraj share a presentation highlighting key points about today's topic, then we will enter the open Q&A portion of the event, and just know that you can submit any questions you may have through the Q&A tool that should be located at the bottom of your screen. Just type in your question and we will respond to each question after the presentation as they come in. You can choose to submit your question, if you wish to do so, anonymously.

Please remember to keep your questions general in nature. If you have any questions about a specific situation involving specific individuals or programs, please email exams@abog.org and one of our staff members will be happy to assist you. At the end of our open Q&A, if anyone has a question that they would like to voice rather than submit it through the Q&A tool, please use

the raise your hand function, also located at the bottom of your screen, and then I will enable your microphone, which will allow you to unmute and then ask your question.

Finally, like all of our other previous webinars, this event is being recorded, and the link to access it will be emailed to everyone on this call, as well as to those who could not attend, so that if you choose to go back and review it or for anyone who wasn't able to join us today, they can have access to the entire session. That is all the housekeeping items I have at this time. On that note, I'm going to stop sharing my screen and I will hand it over to Dr. Shivraj, and she will share her screen.

PS: Thank you so much, Sandra. Welcome everyone. My name is Pooja Shivraj, as Sandra introduced me. I am the Manager of Research and Psychometrics here at ABOG. Psychometrics is everything assessment related that you can think about to do with validity, reliability, and fairness of your assessments and interpreting your assessment data. In thinking about stakeholders that we provide information to, we provide information to you all, and we want to make sure that the information we provide to you is interpretable and is valid in the absolute sense, and so this webinar is on how we have developed these program performance reports, the background of that, and then I will delve into the actual reports themselves.

In thinking about how the information we used to provide before, for those of you that have seen what we used to provide on your portals before, we used to provide a good amount of information prior. We used to provide pass/fail information of all your takers, the number of takers you had, the number of first-time takers, repeat takers, what percent of those takers passed, three-year trends of those takers, the number of three-year takers over the past three years, how many takers took it.

We used to provide that information within a table, but in taking a step back about what assessment data is all about and for you all as Program Directors, how you are going to use that information, how you process that information, and what it is all about. Most programs use that information for program evaluation, for evaluating their program, where their strengths lie, how they can grow, their opportunities for growth, how they can move their program forward. That information that we used to provide was just not enough, and so we thought about that and sort of developed our reports, enhanced our reports and our reporting, taking that into consideration.

Our current program reports provide a more enhanced feedback for purposes of program evaluation and skill improvement. We try to provide feedback at two levels: one at your program level, so you can see how your candidates within your program have performed. We also try to provide information at an aggregate level: how your program performs and how your programs perform in comparison to other programs, candidates in as an aggregate compared to your program across all the other programs.

We also identify your strengths and opportunities or your weaknesses, where you can grow your program in comparison to our blueprint. Also, in looking at the literature in how people process data, we realize tabular data isn't necessarily the best way to go by itself, and so what we have done is also integrated visual displays with the tabular displays we used to provide.

Reports, psychometrically, aren't just thrown together. This is an iterative design process that was used and is common in psychometrics of how school reporting and generally reports that are provided to stakeholders are put together. Internally, I started with a design process of putting it together, but that report might just have jargon in it, statistical jargon, that sort of thing, so it went through an internal subject matter expert review at ABOG to see whether it was interpretable, whether it made sense, that sort of thing, whether it's providing the information that you all need in order to digest it, in order to provide feedback to you all.

Then it went through focus groups with Program Directors like you all. We gathered feedback at conferences from Program Directors, and at each step, we redesigned those reports in order to make them better, in order to enhance those reports to make them essentially what Program Directors want. We also decided to balance wants and needs. We couldn't take in everything that we got, all the feedback that we got. We couldn't possibly take all that information in in order to incorporate that into reports, and the reason being is that we have to balance the psychometrics of it.

A lot of the nitty-gritty information that could be included cannot necessarily be reliable information to provide, so we needed to balance the wants and the needs prior to providing that information. That iterative design process is how we reached the final report that you all now see. That doesn't mean it's necessarily finalized. We still go through this iterative process of design every time we receive stakeholder comments and feedback.

Now in thinking about the reports and the reports that you receive, what is the intended purpose and how have we designed it for you all? We answer four main questions within this report: one is how did your program perform on the exam, on our Qualifying Exam or our Certifying Exam this year, and how did it compare in comparison to candidates in other programs? That's one of the questions that the program performance reports address.

The next question that we want to address is how did your program perform over a three-year period, and how did that compare, that performance on the Qualifying Exam or the Certifying Exam, how did that compare to other programs over that same three-year period? You're doing a your program versus other programs comparison, a one year to one year, a three year to three-year comparison.

The next question we wanted to answer was what are the strengths based on our blueprint? What are the strengths of your program and what are some areas that you can consider possibly improving for your program based on your performance on our exam? That was one of the next questions we try to address in your program performance report.

Then the last question we try to address in the report is your individual candidate's performance, and most of these questions we didn't address prior, so this is an enhancement on the program performance report.

On the first page, or the first tabular display that you see, this is an adaptation of it because it was too small to screenshot. On that first page, the tabular display is a quick snapshot of your data. It provides a here is this current exam, and here is the three-year trend of your exam. It's just a way for you to look at how your program is doing. Looking at what it provides, this is a breakdown of your candidates first time, repeat and a total. This is the 2020 Qualifying Exam, I've just pulled an example program.

If you look across the columns, you can see first time takers, repeat takers, and the total takers. This gives you the number of passing, the number of total takers, and the percent pass. Essentially, if you look at first time takers, you can see that seven pass out of the total seven, which gives you the percentage of zero of one or zero percent, and then if you add across the columns, you can see that, for example, your total takers, you have seven first time takers, one repeat, for a total of eight. That's how you would interpret that table.

The same for the 2018 through 2020. That's your three-year trend. What we have also done in addition is put an asterisk here, which tells you the ACGME requirement. ACGME has a requirement, and we have put that at the bottom of the table. You get a citation for if you're below 80 percent for a pass rate for ABOG, or if you're below the fifth percentile, and it just tells you whether or not you meet that requirement.

The next graph shows you current year's data on where you are in comparison to the other programs. Over here, the y-axis shows you your percent pass rate, yours or the other program, so just a current year pass rate for 2020. The blue tells you where your program is at, and the orange tells you where all the other programs are at. Just in breaking down what that looks like, this is only for first time takers, it's not for all takers, it only includes first time takers, and that you can see in the table below, so this is how the calculation is done and I'll go over the calculation just for this table, because it repeats across the other slides.

If you look at your program and how that 100 percent is calculated, you have the number passing and number of first-time takers, which is calculated to be that hundred percent. For all programs, again, first time takers, it's that 1202 passing of the 1297 first time takers, which leads to that 92.68 for all programs. You're doing that comparison of your current year versus all other programs. But now, most people want to know for my program, if I perform at a hundred percent, for example, am I the only program that performs at 100 percent or most programs there where are all the other programs at for the current year?

The next graph shows you the distribution of where you are at for that current year. If you are at that hundred percent, it shows you where you are at for that hundred percent. The left y-axis shows you the number of programs that are at that percent, the right y-axis shows you the percent of the total programs, and the x-axis at the bottom shows you the pass rate.

Essentially, how you are reading this graph is if you are at that 100 percent pass rate, 169 programs are here, which makes up about 80 percent of the programs. That's how you are interpreting this graph. If you are at 90 percent, three programs are here, which makes up about five percent about that much right, and so that's how you're interpreting this graph. This tells you the number of programs that are in this graph and what your pass rate is, which you can get from the prior graph as well.

This gets to the three-year data. This is a rolling average and what a rolling three-year average is is you have your 2018 through 2020, which is the three year, but you also have it rolling, so 2017 through 19 and a 2016 through 2018. The 2019 is included in this as well, and so it kind of rolls along is what a three-year rolling average is. Again, if you look at the y-axis, it's your three-year pass rate, and you're comparing your blue to your orange, where blue is still your program and the orange is still all programs.

Again, orange, the all programs, is calculated as the graduates of all programs and the first-time takers, so the number passing versus all first-time takers, and that's how it's calculated across all three years. The x over here is the fifth percentile cutoff, and we've just given that to you as a reference point for what ACGME is looking at in terms of the fifth percentile cutoff.

Again, if you think about where you are at, if you look at that 2018 through 2020 graph, and you look at oh I'm at 95.65, how many programs are above me, how many programs are below me? Most people want to know where they are in comparison to the other programs. This gives you that distribution of where exactly you lie in comparison to the other programs, and that's the purpose of this graph. It tells you how many programs are there and where you lie in comparison to the other programs. This is just a distribution of that, again telling you what the frequency is, number of programs which lie in that bar graph, the percent of those total programs, and the pass rate on the x-axis. Over here, if you're looking at a hundred percent of pass rate over a three-year period, 108 programs lie there, which makes up about 45 percent of the programs and that's how you're interpreting that graph.

Next graph we get to is the program performance by content domain. This tells you the min and the max by content area. The content area we focus on is mapped to our blueprint at ABOG. You can see below on the x-axis we have it labeled as OB1, OB2, which goes to G1, G2 and then goes to OP1, OP2 over here, and then C1, C2. If you go to the next chart, the table below, it tells you what OB1, what OB2 is, what C1, C2 is. It states that it's communication, that sort of thing. It tells you, it describes what it is below.

It's too small to fit in the chart. We didn't want little squiggly characters there, so we didn't actually want to elaborate on it in the chart, so we've left it as OB1, OB2, and we played around with this a little bit to make it clean. What the y-axis here shows is the percent correct by content area, so it goes from zero to 100 percent correct. It shows you a min and a max. The blue again being your program and the orange being all programs, and the range, so it can show

you a quick where your program is at in terms of a min and the max, and then we're all programs are at.

Finally, the strengths and the opportunities for your program. Again, the category names match to what the categories were before, so again it starts with the OB1, OB2, and then goes all the way to C3. I've given you your median as opposed to the mean, and the median because if you have outliers in your program, for example if one person has gotten a zero percent correct and the others are all in a regular range, the median tends to be more accurate.

Your median is given here, and all the other programs median is given separately in order for you to compare. The green dots tell you you are generally fine and within half a standard deviation or more of all programs. The red dots tell you you're below half a standard deviation of the other programs, and you might want to pay attention to one or more of those areas. It's just a quick way to look at where you might want to pay attention in terms of a program evaluation for you all.

Finally, the last table tells you your candidate scores. We've provided you with an ABOG ID, your candidate's names, I'll get to this, it's a scale score, and your candidate's exam results, whether it's a pass or a fail. We're providing scale score as opposed to the raw score. We deliver 180 items, and we are not providing just a are they getting 100 items correct out of 180 or 120 items out of 180? We have moved to a scale score now in order to provide fair comparisons across yours, as well as candidates, and it be comparable.

In psychometrics, using scale scores is standard. We provide you with a scale score and below the table, we tell you what the passing scale score is for candidates. We also provide candidates with this information, so they know what their scale score is and what the passing scale score is for them as well. This information should be provided to you in the table, however, if this is blank, the candidate did not give us permission to release their scale score.

Now there are some differences between the Qualifying Exam and the Certifying Exam reports. As I told you, it's an iterative process of design. We did start off with the Qualifying Exam reports when we initially started our design process, and the Certifying Exam currently as it stands does not have the last couple of things I described, which is the program performance by content domain.

We currently do not have the strengths and opportunities for improvement, and the Certifying Exam does not have the scale scores. We only provide the pass/fail information on the last page. However, these are things we are working on for the next iteration of the report, and enhancements will be added to the next iteration of these reports in order to provide you with more information that you can work off of.

That being said, this next cycle of the Certifying Exam, the 2020-2021 Specialty Certifying Exam program performance reports, given that it will be a multiple-choice exam similar to the Qualifying Exam, the program performance reports will look very similar to what you saw just

now with the Qualifying Exam, given that we can provide more information, more detailed information. If you think about the Certifying Exam and the structure of the Certifying Exam, candidates take their exam in an oral format. They take their assessment on case lists and things like that, so in conceptualizing those reports, they are just giving information just in a different format, and so conceptualizing it is just different than providing information on the Qualifying Exam reports. Given that the 2020-2021 exam will be similar in format to the Qualifying Exam, your program performance reports right now will look similar, but in the next cycle of the 2021-2022 CE report, we will be implementing those new changes.

I do want to talk about a couple of limitations and interpretations for these reports. They are on your first page of your report, but as a psychometrician/statistician, I do want to ensure that they are interpreted appropriately. Programs with fewer candidates will have less stable results, lesser precision on those results. When I say fewer candidates, smaller residency programs, specialty, subspecialty fellowships, those are included. Programs that have a variable number of candidates taking both ABOG and ABOG exams each year will also be impacted by this, and so just keeping in mind being cautious about the way you interpret these results. You just need to keep that in mind while going through that.

The next thing I wanted to make you all aware of is that the strengths and opportunities graph, as well as the program performance graph. The program performance graph you notice the mins and the max. There might be outliers, so that's something to keep in mind when thinking about looking at your places to evaluate your program. But thinking about the strengths and opportunities and where those reds and the greens lie, implementing any sort of program change will take a while to see that three-year trend change, so don't expect to see a change immediately in your three-year trend as you implement change. Of course, implementation of any sort of program change may not be reflected in candidate scores, but providing that information of where your program needs those opportunities for growth, and has those strengths is important as well, and I think it maps to the blueprint pretty well.

That's all I have in terms of the report and the descriptions. Thank you so much for attending, and I will take any questions at this point, along with Dr. Larsen.

Our GME office requires five years of board passage data for our program APE. Where can we find this information? Previously, we were able to pull it from ABOG.

WL: That's not obviously currently in the program report, but as Pooja said, this is a process that as we get feedback, we try to change how we provide the data. If that's something that programs need us to provide, then we'll add that as a future enhancement to the report, but for an individual program that needs it now, please just email us at exams@abog.org with the

information about your program, and we can certainly calculate that for you and get that back to you.

Were there any areas of general residency program performance below what was expected for the 2020 grads whose training might have been disrupted by the COVID-19 pandemic?

PS: For that, we are actually looking into that right now. We don't have an answer yet, unfortunately, but we are looking actively looking into it right now.

Does the Program Director have access? It used to be the PC or Program Manager who did?

PS: Yes, I believe the Program Director has access now to those reports.

Can you explain the programs with zero percent pass rates?

PS: Yes. The zero percent pass rates are the people that have just had zero people pass the exam out of either they've had one taker take it and zero people pass it, so they've just had a small number of people take it and pass it.

WL: In other words, it could be a really small residency where not everyone took the exam this year, or it could be a previous AOA-certified residency that's now under ACGME and one of their graduates chose to take the ABOG exam but didn't pass.

How will the ACGME-managed data from programs with some AOBOG and some ABOG candidates?

WL: I think that's a question that you'll have to ask ACGME. We provide obviously the data on our pass, the passing program data to ACGME, but they will make the decision on how they'll manage programs like that.

Is there useful performance data from repeat takers?

PS: We don't provide information about repeat takers because repeat takers tend to skew the information, the data, that we provide. That's why we provide repeat taker information on the very first page of the program performance reports just to show you what that looks like, but in terms of your three-year trends and where you can improve your program and that sort of

thing, we only include first-time takers because repeat takers don't tend to be part of your program and skew your data.