BHW (Health Workforce)

Health Resources and Services Administration

Student-Level Longitudinal Tracking for PCTE Awardees

3/29/2017

Essex, Alyson (HRSA)

Submitted by:Simon, Nolan (HRSA) [C], *Adobe Connect Team*

Event: Student-Level Longitudinal Tracking for PCTE Awardees

Date: 3/29/2017

Event Coordinator: Essex, Alyson (HRSA)

Adobe Connect License: Seminar (<1,500 participants)

Unique Users: 102 unique users

Audio: Universal Voice/ Conference Bridge

Start and End Time: 2:00 – 3:30 PM EST.

Duration: 90 minutes

URL: https://hrsaseminar.adobeconnect.com/tracking_for_pcte/

Problems Encountered with Adobe Connect Pro

No Problems Encountered

Recording

https://hrsaseminar.adobeconnect.com/p2hdtddbo7b/

Attendees

Meghan Chambers
9833734
Abayomi Salako
Alai Tan
Allen Perkins
Alyson Essex
Alyssa Adams
Amy Bethge
Amy Cunningham
Amy Vega
Ana Marin
Anita Franklin
Ann Stoutenburg
Anne Patterson
Anthony A
Ariel Ausmer
BEP
Beth Hribar
BGJ
Bonnie Vest
Brandi Leach
Brooke Gillespie
Candice Chen
Caption Colorado
Carla Terry
Carmen Ingram-Thorpe
Chenelle
Christian Neal
Christine Morgan
Christine Schudel
Colleen Payton
Craig Stevens
Cryst
Daniel
Darby Ford

David Sacks
Deborah Center - Colorado
Center for Nursing
Excellence
Denise McGuigan
El Capiton Botkin
Erin Griffi
Geri
Gina Flack
Harry Mazurek
Heather Miselis
James Madison University
Jamie McCauley
JBS-IT
Joann Petrini
Joy Bowen
Julie DiGregorio
K Hills
Kate Holeman
Katherine Mott
Kathleen Curtin
Kelly Budke McCarthy
Keren Wick
Kristin
Kristin Baughman
Laurene Enns
Linda Dale
Lyle
M. Davis
Mallory Johnson
Manu Singh
Maria Montoro Edwards
Maria P
Mary Crane
Mary Pileggi

Mary Springston
Melanie Caldwell
Michael Steigerwald
Michela Bull
Nadiya Pavlishyn
Natalie Truesdell
Nolan Simon
Pat Matthews-Juarez
Patricia Dieter
Patricia Hogue
Perri Morgan
Peter DiMilia
Raphael Gaeta
Ruth Dufresne
Sandra Banas
Sandro
Sarah
Sonji Miller
Sridhar Chilimuri
Stacey Gardner-Buckshaw
Stephany Thompson
Susan Moyer
Svetlana Cicale
T. Riley
Taylor Grimm
Tiffinie Zellars
Toho Soma
Tom Bik
Tracey Smith
Union Hospital - HR
University of Cincinnati
Valentín García
William Moran
Zuay Thompson

Chat History

N/A

Polls

N/A

Q&A

1. Sandra Banas: Will a copy of the PowerPoint be available after the webinar?

2. David Sacks: Can you discuss any "hiccups" you overcame as you worked on database over the past several years?

3. Maria P: Irene Sandvold has a question: Can you clarify what one big database for all of our programs would look like?

4. Maria P: Another question: What pre-graduation activities have you found have been predictive of post-graduate leadership positions?

5. Christine Morgan: For IRB consent, how often are you asking for consent, every survey?

6. Alyssa Adams: Great presentation. The article mentioned is linked in the webinar announcement that was emailed out

7. Pat Matthews-Juarez: How long did it take you to create the data base?

8. Maria P: Do you also have a high response rate to the SES [socio-economic status] questions like clinical income as well as the sanctions questions?

9. Ruth Dufresne: You mentioned that Duke started this database 4-5 years ago and you've had institutional support. Can you say anything about the resources needed to get a database like this started? And how many records are in the database now? Thanks.

10. Maria P: Irene Sandvold: How do you specifically increase your response rate; for example, what types of incentives or funding do you provide to achieve this?

11. Pat Matthews-Juarez: Thanks

12. Deborah Center – Colorado Center for Nursing Excellence: Thank you

Transcript

Primary Care Training and Enhancement (PCTE) Program Evaluation Technical Assistance (TA) Webinar Series

Webinar #2 Title: Creating a Trainee-Level Longitudinal Education Database: Conceptual and Methodological Considerations
Date: Wednesday, March 29, 2017 at 2:00pm ET – 3:00pm ET
Meeting Details:
- **URL:** https://hrsaseminar.adobeconnect.com/tracking_for_pcte
- **Conference Number:** 888-942-9693
- **Participant passcode:** 9833734

[Please standby for real time captions] Thank you for standing by. The conference will begin in a few moments. We ask that you continue to hold.

>> Will come. At this time lines are in a listen only mode for the presentation. The conference is being recorded if you object to disconnect at this time. The conference features a Q&A session if you want to ask a question please press*one. Altering the conference over to [indiscernible] begin when ready. >> I think we should begin [by asking] how many folks do we have the call right now? We have about 33 standing by in your line is open.

Thank you, so much and good afternoon, on behalf of HRSA I want to welcome you to participate in the second webinar series that we have for our PCTE grantees. I am the contracting officer's representative for the Primary Care Training and Enhancement, or as we all know, PCTE programs, within the National Center for Health Workforce Analysis in the Bureau of Health Workforce here at HRSA. The overarching purpose of PCTE is really to strengthen the primary care workforce by supporting enhanced training for primary care clinicians, teachers, and researchers, and this program focuses on supporting innovative training in three primary ways. The first is to encourage high-quality care – primary care – in underserved areas. Another is to enhance diversity within the workforce, and the third is to produce clinicians who will practice in, and lead, transforming health care delivery models. So again, we want to thank you for joining us today and we are very pleased to offer this series, which is specifically designed

for our PCTE awardees, but we know that many our – the topics that we'll be focusing on are applicable to other HRSA grantees as well and we're hoping that these evaluation resources are really going to help increase your capacity to develop and conduct your own evaluations. These educational webinars will focus on the application of a particular evaluation topic area and concepts that are relevant to the work that you're doing as one of our PCTE awardees. As a kickoff to our second webinar series, we're going to be discussing conceptual and methodological issues in creating a trainee-level mod (?) to your educational database. Prior to turning the conference over to Manu Singh, who's one of our colleagues, I just would like to quickly go over our agenda for today. We'll transition first to some welcoming [and] housekeeping by our contactor for the PCTE contract and then we will have presentations by two leaders in the field – Dr. Perri Morgan and Dr. Brandi Leach, both of the Duke University [School of] Medicine. And then finally we will close out with 10 minutes of a Q&A session, so please feel free, if you have questions during the presentation, to jot them down and we will have an opportunity to address them at the end of the call. Manu? I'm going to go ahead and turn it over to you.

Thank you so much. I just wanted to check in – Nolan, I am just having a hard time forwarding the slides and advancing the slides on here. I just want to check to make sure that is corrected before we move forward, so I can make sure that we should be on...

Are you seeing the directional arrows on the bottom left-hand corner of the presentation?

I do.

Okay, can you just forward them on for me at this point? Maybe others will have better luck.

Sure.

Thank you. The next slide? Oh, actually it's because we need to – this is the lobby slide – that's what's going on here. Thank you so much. Perfect. Okay, great. Perfect. So, I did want to just take care of a few housekeeping items, as you can see. One of the pieces is, if there are any technical difficulties on your end, as the participants listening to this, please do use the Q&A

box. You can see it says "Q&A." If you can go ahead and type in any questions you have related to any logistics, that would be great, and we can attend to that as soon as possible. Any questions related to the speaker or presentations, please hold those to the very end and then the operator will announce the time at which when you can ask those questions and we will address as they come in, and the operator will read those questions aloud and address them to the speaker that they're most relevant for.

A couple other things I do want to mention is that today's meeting material will also be available on the HRSA webpage, and we'll notify you when those are available, including the slides as well as the whole webinar. And the other piece is, as you can see on your screen – and I'm just looking, and I don't – is the poll there now? Nolan? Shortly you should be able to see right underneath the Q&A box, there's going to be a little poll. I know at the beginning when you all signed in, many of you were asked to provide your name, your discipline, organization and email address. If you did not provide that because they stopped collecting that information, could you please complete this poll and just type in that information for us, so that we have everyone's information and can get you any resources and materials that you might need after this webinar.

Let me just make sure there is nothing else remaining here. I think that's it. I'm going to go ahead now and turn it over to a Craig Stevens, a member of the JSI/JBS Evaluation Technical Assistance Team, who will provide some context on the purpose and goals of the webinar. Craig?

Thanks, Manu. So, as Manu mentioned, I'm just going to spend a couple minutes framing, you know, why are we having this discussion regarding the educational longitudinal databases. So, as you all very well know, there are specific areas of interest for PCTE by HRSA and these are represented in the RFA's and your evaluation plans, including disposition in primary care and underserved settings, as well as how well your trainings addressed – essentially the quadruple aim – the access, quality, cost, satisfaction, or resiliency. So, at the core of this, what it really is

requiring is taking a look at tracking trainees' post-program completion to collect some of this data, so we're going to talk about that briefly.

One of the key messages that HRSA and our team has talked about is the balance of rigor and resources in doing this work. Certainly, we are looking for high rigor and adding to the body of knowledge but realizing that that's not necessarily able to be accomplished for everyone because of the potential resources. So, as we go through this, thinking about what are the goals of your program, what's going to be valuable and feasible, is definitely something that should be taken into consideration. Taking a look at tracking trainees post-program completion, as you very well know, sometimes it's difficult to track trainees and it can be resource intensive. And what we have spoken with some awardees about, through interviews, site visits, and discussing their evaluation plans is to consider incremental approaches or targeting certain trainee cohorts. So, some folks have talked about targeting trainees that come back to their state to practice, working with state licensing boards. Others have talked about targeting trainees that land back in their health systems where measuring quality, for example, is actually feasible. So, consider those incremental approaches in your work.

Collecting data -- outcome data – relevant to the quadruple aims, depending on who you're tracking can be difficult. So, if you're asking about EHR data for a trainee that's outside of your system, that may be very difficult to obtain. So, consider what are some limited and programmatically useful proxy measures if the outcome measures are not able to be obtained, and I'm going to give a couple of examples of that. So, continuing on, collecting data, a couple examples that could be collected for proxy measures is whether a trainee, post their program completion, is participating or leading in QI [Quality Improvement] initiatives – as you can imagine, [this is] a proxy measure to quality. Whether they are participating as a preceptor, faculty, whether their writing publications or engaged in research – a proxy measure to whether or not they're satisfied or whether there is resiliency. Whether they're participating in healthcare reform such as shared savings, alternative payment models, that are driving them toward the balance of quality and cost. And then whether or not they are participating in a

leadership or management role in clinical practice – so these might be good proxy measures as you're building your capacity to do more rigorous work.

Then the last is the analysis. So again, you may be working with small numbers if you're targeting a small cohort, but HRSA is interested in – and your program should consider – what insight can be gained for the proof of the program improvement versus if it's moving toward rigorous research. So again, rigorous research would be wonderful in adding to the body of knowledge in this area, but balance the rigor and resources, so at least there is insight, if not moving towards something more rigorous.

A couple of examples of resources that other awardees have relied on to help with this work. Some folks are working with their alumni affairs and their development offices because they actually do an excellent job of tracking people post their program completion. A number of Area Health Education Center [AHEC] programs, which I believe are in every state, do tracking of healthcare professionals, because this is also within the scope of their evaluation. And in fact, some AHECs – the central New York AHEC -- has developed programs that are based on social media platforms that connect trainees together as well as provide opportunities for them to be tracked over time and to push-out questions or larger scale surveys. We have some folks, as I mentioned before, working with state licensing databases. I'll speak of my home state of Vermont, actually does a physician licensing survey and then an advanced practice professional survey that asks a number of detailed questions about their practice, their setting, [and] their specialty, and those may be accessible to you. There may be commercial tracking databases that are available, as well as national resources such as the NPI [National Provider Identifier] and the AMA [American Medical Association] Physician Masterfile database.

So, with that, I'm going to turn this back over to Dr. Essex and as we mentioned, we'll hold questions until the end.

Thanks, Craig, so much for providing that additional contextual information. We just wanted to quickly ask before we move forward if anyone has any questions about what Craig shared or about or about the strategic initiative goals and objectives of the PCTE program and HRSA and

BHW as a whole. My contact information is listed on the slide. Feel free to reach out to me by telephone or email at any time. Now, without further ado, I want to take the time to introduce our two guest speakers for today, Dr. Perri Morgan and Dr. Brandi Leach. Dr. Morgan is a health services researcher, and she focuses on physician assistants [PA]and nurse practitioner [NP]s in the health workforce, and on outcomes associated with their use in different roles and settings. As Director of Research in the Duke PA Division she has led the development of the PA Research section, and also after practicing for 25 years as a PA she has knowledge of the PA profession from the perspective of clinicians, which is incredibly important. Her research addresses methodological problems of data sources for use in research on PAs and NPs, the effective PA use on quality of care and health resource utilization, and the use and roles of PAs and NPs in various settings. Dr. Morgan received her PA training at Emory University and her Master's and doctoral degrees from the University of Wisconsin at Madison.

Our second presenter is Brandi Leach she is a social -- sociologist and health services researcher and investigates the changing healthcare system within the US, with a specific emphasis on workforce issues, key-based care, the role of PAs and NPs within the healthcare system, and health professions education. Dr. Leach is currently employed as a research analyst in the Duke PA division where she manages the education research database, and that's a longitudinal database tracking system for PAs from training through their early career. She received her doctoral degree from North Carolina State University. Now I will turn it over to our first speaker Dr. Perri Morgan. Dr. Morgan, thank you for presenting today.

And Kyle, if you could go ahead and put Dr. Morgan's slides up that that would be wonderful. Thank you.

Hello, everyone. It's a pleasure to be with you today and to share with you a little bit about what we have done with creating our longitudinal database. We started this about four or five years ago now. It was not developed as part of our HRSA grant, although we do in our program have HRSA primary care training expansion grants. It was developed for our own research and evaluation processes, but I think the folks at HRSA thought that it could be adapted to be used

for HRSA grant evaluations and that's why they invited us. So, we hope you'll find this to be useful.

So, we're going to talk about why you might consider using a longitudinal education database, show you some examples of studies that have been done with databases like this, and then summarize our conceptual approach, some of our data sources, and talk to you about some of our processes.

So, why would we want to do a longitudinal database? Obviously, as educators we are interested first and foremost in the development of our students over time, thereby a longitudinal process. And so, we really need to ideally, I think, follow individuals over time. I'm going to give you an example next and I hope you won't be insulted by the simplistic level of this example, but I find it useful to start with this very basic, straightforward example. This is how I explain the – when we get our students – we ask our students for informed consent to participate in the database, these are exactly the slides we show them, so will start with that.

Imagine you have two PA students entering a PA school and you ask them you're interested in how likely they are to go into practice as a surgical PA, and you ask them to rate that likelihood on a scale of 1 to 10. If student Anne says, "No way. I don't want to go into surgery" and answers with a one. Your student Sue says "Absolutely, that's what I came here for and I'm going to be a surgical PA," and she answers a 10. At graduation you ask them the same question and they flipped positions. Well, I expect you can see where I'm going with this.

If we analyze this cross-sectionally and take the mean of our first-year group and the mean of our graduating group, we see that both times the mean was 5.5 and there was no change in the students' attitudes toward surgery would be our conclusion. If we analyze it longitudinally however, we can see that both students changed by nine points and the mean change was nine points and we conclude that student attitudes about working in surgery change a lot during the course of their PA education. In order to be able to do this longitudinal analysis we have to be able to link individual student answers – in this case, their first response to their latter response.

So, we chose a student-level longitudinal approach for our database for several reasons. We wanted to avoid the potential fallacy that I just showed you – of equating group change with individual change. We also thought that this would make it possible for us to identify students with certain characteristics – so, certain subsets of students that we might be interested in for particular research questions – and then contact those students and ask them if they were willing to participate in these, sort of, sub-projects. We also thought, you know we collect all of our data, as I expect you do as well, at the individual level anyway, so why not use all of that extra information that you gain by keeping track of which individuals answered which way at which time.

So, in deciding to set up one of these databases it's essential to think about whether or not you are going to use the database for evaluation or for research or for both. So, I think it's useful to take a moment to compare the two purposes. So, for research, we are interested in producing generalizable knowledge and also perhaps disseminating that knowledge. We use scientific methods and we typically have to get IRB [Institutional Review Board] or human subjects review approval. For evaluation, on the other hand, the intent is to improve a specific program and the findings are expected to directly impact that program and to identify potential improvements in only that program, and it's geared toward making decisions at the program level. Often this kind of evaluation does not require human subjects review. So, I think for the purposes of what you folks are doing with evaluating a HRSA training grant, you could choose to do either research or evaluation.

So, let's think about why you might want to do each of those. If you are interested in doing this solely for evaluation, a longitudinal database might help you organize all of your data. We've realized for years here, and at other PA programs where I worked, we have all this very, very rich information, but it's located in different places all around our educational program, and this educational research database allowed us to put it all in one place. You can use that information for program improvement and you might be able to analyze issues that are specific to your students or to your program. So, for example, does a specific admissions factor predict the specific problem you see amongst your students, or does a specific intervention in your

education program work better for a particular type of student in your setting? Finally, you might choose to do an evaluation database because you don't want to deal with it human subjects review and informed consent process, but we think this is a weak excuse.

On the other hand, why would you want to create a longitudinal database for research purposes? If you do this, we can do all the things on the previous slide and in addition we can share our findings with other programs and the education community, be it presentations or publications. We can help our faculty colleagues who need to produce research, to have a data source that they can use to answer interesting and important questions. Also, we think that we can leverage the data that we have already collected to enrich new projects.

 So, for example, maybe a faculty member might want to look at or consider outcomes of students who participated in international rotations, for example. We would already have all those students' demographic and background data, so the investigator who wanted to study international rotations could do a relatively short survey, for example, and then link it to all the other variables that we had about those students. Therefore, the surveys can be perhaps shorter, which hopefully could reduce survey fatigue among our students. Finally, if you are set up to do research and produce generalizable results, you might be able to combine your program data with that of other institutions in the future to do research that spans institutions and/or expands professions.

So, think for a moment about when evaluation – if you decided to go the evaluation rather than the research route for your database – when does it not require human subjects review? I just want to emphasize – I heard this mentioned a little bit before – but emphasize these three points. It doesn't require human subjects review when it does not involve nonstandard interventions, so it's double negative – so when you are doing fairly typical interventions. The intent is to only provide information for and about the setting in which it's conducted and when it's part of your standard operating procedures.

As I mentioned earlier we decided to go with a research database, so therefore we did go to human subjects' review, and our experience was actually quite positive with that. The way we

approached it was we created one IRB proposal that was about creating the database; and so, the students consent to be in the database, but not in any particular study, at the time the database is done. So, the IRB number refers to the database and then if we decide to do a project, for example maybe one looking at students who participated in international rotations, the faculty member who's interested in doing that project submits a separate IRB proposal that references our database as the data source. Once we got the additional IRB proposal approved, most of our amendments – and there have been many because every time we amend a survey or change a question we send it back through – but these are expedited usually with a two- to three- day turnaround.

So, let's think about some examples of education research using longitudinal database. The Jefferson Medical School started a longitudinal database of their students back in 1970 and they might still be going on; it's gone on for a very long time. They've published at least 150 articles out of this database, and here are a couple of examples. This was one published in 2005, which looked at disciplinary actions by medical boards and then compared that to prior behavior problems in medical school. In 2007, there was another article listed here where physician scores on the National Clinical Skills Examination were predictors of complaints to medical regulatory authorities.

We have a few projects starting out now similarly in our program. For example, we are looking at how students' expectations of salary compare by gender of students entering the program and graduating from the program, hoping that this will give us some insight into how to correct actual salary disparities among graduates.

So, our current database is this permanent database that contains extensive longitudinal student-level data from the PA program. I'm going to go through many of the data elements here in a minute or two. We start data collection before the students ever enter. On our supplemental admissions application, there is a short consent that [states] whether or not they allow their admissions data to be included in our database. Therefore, because we started so early we're able to support research on PA selection, on the training process, and on practice.

So, conceptually very big high-level picture, we collect data at the pre-PA school time point, all during PA school, and post-PA school, and then we link that data by individual across time to create what is essentially a very, very large spreadsheet on each student.

So, for our pre-PA data sources these are some of the – this is some of the information that we collect. Much of it from the admissions application, so we have the essays – some we could do, if we wanted, we could do research using qualitative analysis of admission essays, we have their grades from college, their veteran status, we have evaluations of their interviews when they came to interview, what their admissions committee said about them, what their healthcare experience was before they entered PA school. We also have information that was collected at a survey when they arrive at our program, which involves how much debt they have coming into PA school and some important socioeconomic status indicators of their family of origin or their financial situations just before they start PA school. We also include a number of psychosocial factors that measure at the point at which they join us here, including their family structure, whether they have a partner, whether they have children – dependent children, as well as some measures of stress. Most of this data, as I've said, came from the admissions process and our new student survey.

For information during PA school, we basically include everything that we know about these students is kept in records; so, all of their grades, all of their clinical rotations they participated in, what their scores were from their preceptor evaluations, if we have any particular programs to emphasize particular areas, such as the primary care track, that would be noted here. If there was a HRSA program, for example, that only affected some of our students, that would be noted here. We also have, of course, whether they passed their national boards. We also, in addition to that, sort of, more straight, academic information, we collect information how much they participated in volunteer activities throughout our program and whether they obtained any awards, whether they had held leadership positions in the student society, etc. Then, we continue to survey them about any changes in their family structure. The stress survey is given at three points – at matriculation, mid-point, and graduation – as well as information on social support, major life events and personality traits is one of the enduring

characteristics that we are interested in, so we collect that only once because we don't expect it to change; but many of these other facets we collect repeatedly, often at entry, at midpoint, and at graduation.

Then, after PA school we continue to collect information and add it to this database – to this big old spreadsheet – which includes whether the student is practicing, what kind of position they took. Because we're fairly new at this, we have only done the one survey, which is the survey that is six months after graduation. We also plan to do another survey when they have been out in practice for two years. This survey includes things such as what specialty did they practice in/take a job in, where it is geographically, what sort of setting they are practicing in, what portion of their patients have different payment forms, etc. We also ask about income and we ask about whether they are participating in loan repayment. Also, things such as job satisfaction, and stress and burnout will be addressed in the future. We did not ask that six months out, because we felt it was a little too early. We hope they're not burned out yet.

Another section we have not yet collected outcomes in, but that we think is one of the real major areas of promise for this approach, is to be able to ultimately link each individual who went through our program with their NPI numbers, so that we could use electronic health records and other data sources to evaluate their outcomes in terms of cost of care, quality of care, and providing access to care.

So, looking back at those last three slides, when I talked about the information that we gathered pre-PA school, during PA school, and after PA school, you can see that this slide follows. So, these are the kinds of questions we could ask using this database. For example, "What characteristics predict admission," that would be using pre-PA school data mostly. "What PA program experiences are associated with postgraduate leadership positions?" We'd be using the PA school and the post-PA school [data], etc. Then, finally, "What PA program experiences are associated with delivery of high-quality care?" One of the most important questions we could ever ask and answer, and that would be using PA school and post-PA school data.

So, we have not used this for evaluation of a PCTE grant, but we are going to take a stab here at trying to transfer what we have learned on to what you might be needing it for and thinking about some of the types of outcomes that you might be able to collect information on. And Craig earlier certainly mentioned some of them. I think the next slide will actually help me to talk about these. So, this is what you have already seen on the left part of this slide and then we can think about potential variables that we could add. So, for example, frequently with a HRSA grant, they're very interested in being able to report the rate or percent of program graduates who are practicing in primary care, so that may come from a post-graduation survey that asked for practice specialty, like the one I just mentioned. For quality of care, one way to do this would be to use the provider ID to link to the claims data, as I mentioned earlier; and, also to look for regulatory board actions, such as the Jefferson Medical School examples I mentioned earlier. Then patient services that are provided by program graduates that could also be evaluated using claims data. Brandi and I had thought that we would take a break here, but I think the webinar is set up to run better if we go ahead and finish. So now my colleague, Brandi Leach, is going to talk to you about some of the nuts and bolts of how she manages our database.

Thank you, Perri. So, before we began, I would just like to reiterate that our database was designed explicitly for research purposes, so it contains some specific – research-specific components, such as an informed consent process, that won't necessarily apply to evaluation-only databases, and I will try to make that clear as we go through the nuts and bolts.

To begin, there are a couple types of data you might not include in the research or a longitudinal database. So first, you wouldn't typically include information that requires anonymity to collect, because you wouldn't be able to match it to an individual student or trainee. So, in our program this includes student course and program evaluations, where we feel that guarantee of anonymity is essential for receiving candid feedback from our students, and we just don't want to jeopardize that. The second type of data you might not include in a research database are data that are required for programmatic administrative purposes and this is because participation in research databases must be voluntary, and so you do not want

to make any collection of data that you absolutely need to collect, dependent upon your students' or trainees' willingness to participate in the database. But there are ways around this. So, for example, in our program we actually collect a fair amount of this type of data, as Perri mentioned, so we collect a lot of course grades and other academic data, but we only include these data for students who have consented to participate in the database, and we let them know that we are doing this when they sign the informed consent form.

Next, I'm going to go over some of the practical issues that will hopefully walk you through the essential steps of the process. So, let's start by talking about student participation and retention. In all likelihood, the students that you wish to collect data on are will be very busy people with many competing demands on their time. So, you will need to sell your project to your audience. One way to do this is to show that your program's leadership team supports the work of the database. This could be in the form of an email from someone in a leadership position, or in our department the Program Director mentions her support of the database during the students' orientation session. Another way we sell the database to our students is by emphasizing the importance of their individual contributions and by explaining how the results of the research has the potential to impact their profession or the experiences of future students. Then, of course, we also like to offer incentives for participation – sometimes it's snacks, sometimes it's a raffle prize; and then, because we are discussing longitudinal databases, it's not enough just to get your students to take a single survey. You need them to stay engaged and to keep taking surveys, even after they leave your program, when they get more busy. You can use this same approach for later data collection efforts that you used when you initially introduced the database, such as offering incentives, but then there are a couple of additional things you can do. So, we like to share with our students any results of our research or your evaluation – anything interesting you have done with the database, so that they can see in very concrete terms what their contributions have made possible. And then we've also found it helpful to identify a student before they leave the program who will act as a champion for the database. And then we ask the champion to promote any of the post-program/post-graduation surveys by reaching out to their fellow alumni through social media or other informal means.

Then, if you are doing a research database, as with all human subject's research, it's essential to complete an informed consent process. This would include reviewing key aspects of the database design with your students, including any potential risks or benefits of participation. We do this in a short presentation to our new students during their orientation week, and after the presentation we give them an electronic copy of the consent form and give them at least a full day to review the form and consider participation. It's very important that students don't feel pressured or coerced into participating, and that your program protects their right to confidentiality by ensuring that faculty members or others don't know who chooses to participate. So, we don't want anyone to feel that they may be viewed differently by their instructor, for example, based on their decision to participate or not participate, and so when we collect the consent forms there aren't any faculty present in the room, in order to observe who is choosing to participate. As Perri mentioned, in order to collect application data on students even those who didn't get into the program, we include a little informed consent statement on the application. We have had a really good success rate with this; last year 93% of applicants agreed to participate in the database.

Once again, privacy is very important, so you need to protect participants' privacy at every step of the process. So, specifically, we protect privacy by giving every student a numeric identifier that we store separately from the database, and then student names or other identifiers, like email addresses or things like that, are never entered into the database. The identifier itself is only available to a limited number of people. That said, it might still be possible to identify students based on a unique combination of characteristics, such as if you know there is only one 35-year-old male who has three kids; so, we guard against that by not releasing the full data set to faculty or other researchers or people who wish to use it. Of course, if we publish anything off of it, we only publish aggregate statistics. Then, finally, we let students and faculty know any efforts to identify students in the database would be a breach of both institutional and federal guidelines for conducting human subjects research.

So, storing the data itself – there are several things to consider about the digital storage of your data. First, it should be stored in a secure space where access is limited to only those people

actively involved in the database project. People who just want to use the data for a single project, such as evaluating a specific program or a small project, such as the example of experiences of international students, shouldn't be given access to the full database, but should instead be provided with a limited, de-identified database.

Then, the second consideration with data storage is just which software program you're going to use to collect all of the data. You're going to be getting data from lots of different sources, so choosing a database program to bring it all together is important; and there really isn't a single program that's better than others, but there are some things you might consider. First, you should consider the level of institutional support that you have available to you. If your department or office is already using a program, then they're more likely to be able to provide you with technical support when things go wrong. Then, you should also consider any specific needs based on your institutional environment and the preferences of the main database users. So, in our program, we had to meet some very strict privacy and data security measures, such as the ability to set specific user rights and also maintaining an audit trail of any changes made to the database. This is in part because we were doing a research database. Then you also just want to consider how you will interact with the database. If you're going to have multiple users, you'll want to look for a web-based or a server-based program. And then you always need to think about both the compatibility of file formats and, of course, what your budget is.

These are just examples of database software options for you to look at, at your leisure. I'm sure there are others, because new ones come out all the time. We ended up ultimately choosing REDCap [Research Electronic Data Capture] for our project, and that was in part because of the strong institutional support for it here at Duke. Your choice may be different based on your own needs.

Then, after you choose your software and set up your database, it's time to actually add the data into it and there's just a couple things to keep in mind while you're doing that. Of course, as with all data the format matters; so hopefully you thought about this step when you chose your database software and insured that there was data compatibility. If not, I will just mention

that there is a great program called Stat Transfer that helps you reformat data from one format to another; it's been a lifesaver for me, as far as getting data into the database. Then also, depending on your data sources, there might also be some IRB issues or data use agreements that you need to consider before you move data into your database. So, this may especially be an issue if you are using data collected by third parties; so always make sure you have the right to use the data in your database and that the participants whose information you're adding to any – particularly research database, but really any database – has consented to the inclusion of their information. Ideally, you'll find a way to note any of these IRB or data-use restrictions of the database itself, so that moving forward you will be able to keep track of them. Then, last, and this is also common to projects, before you add new data to your database you'll most likely need to clean it, so especially when you're bringing in data from a variety of sources, you are not going to have control over the format of a lot of the variables and particularly with a longitudinal database, the specific survey questions or variables may – the format of them may change over time, and so you need to be vigilant to look for any changes in answer categories or question formats, and to make sure to harmonize the data, so that it is compatible across years. That is one of the unique challenges of doing a longitudinal database.

So, it's our hope that many different medical education programs will develop their own research databases that could fuel future collaborations across institutions and we would love to be able to combine data from the ERD with that of other programs. We hope you found this webinar useful and are maybe even feeling inspired to go create your own longitudinal research database. So, we'll go ahead and open it up for questions and thank you.

Thank you. If you would like to ask a question on the phones, please press star one on your keypad, unmute your phone and record your name when prompted. To withdraw your question, you may press star two. Once again, if you'd like to ask a question, please press star one and record your name when prompted. One moment for the incoming questions.

We do have our first question in queue. One moment for the name please.

Our first question is from Carla Terry. Your line is open.

Yes, I was checking to see if the slides would be available from today – if they'll be sent out to us.

Hi, this is Alyson form HRSA. We're currently in the process of putting together a site where the PCTE program, within the grants page on their Division of Medicine site within the Bureau of Health Workforce website. Once we are able to get clearance, there is a – I don't want to say it's a lengthy process – but we need to go through clearance through the Office of Communications. We're currently in the process of drafting the framework and getting everything together for that site. Once that site is up and running, we will be providing all of the audio, PowerPoint slides, any articles from presenters, as well as other additional evaluation toolkit resources that you might find helpful. Once the site goes live we will send out an announcement through your POs [Project Officers] to let you know the information is available. But unfortunately, right now, until we go through the formal clearance process, we are not permitted to release the slides from today's presentation. But do know that we will get them to you as soon as possible.

If what you're interested in has to do with the slides that Brandi and I showed, we do have a publication based on these that was in the Journal of Physician Assistant Education. If you e-mail me, I would be happy to send to you, or you could probably just find it online. It was 2015, and I'm the first author, Perri Morgan.

Hey Perri, this is Alyson. If you would be willing to send me that citation, or I can look it up on PubMed, then perhaps we can ask our POs to disseminate that information to the grantees, so that they could have that as a reference.

Sure. I would be happy to.

Great, thank you so much. Do we have any other questions?

Currently showing no other questions in queue.

Great, okay let's just give it a quick moment. Just so folks know, our next seminar will be in two months – it's March now – so that will be sometime in the end of May. And like we have done with the first two seminars, we will send out an announcement highlighting the topic for that seminar, a highlight a brief description of the presenter, as well as the logistical information in terms of how to connect with Adobe Connect. But, in the interim, if you have any questions or concerns feel free to reach out to me and I would be more than happy to answer your question or find the appropriate person and get the information that you need. Do we have any other questions at this time?

Excuse me. It looked like there were a few questions that came in – this is Manu – there were a few questions that came in to the Q&A box. I'd like to just go ahead and read those if that's okay. From David Sacks, he asked "Can you discuss any hiccups you overcame as you worked on the database?"

Can you repeat that one more time, please?

He said, "Can you discuss any hiccups you overcame as you worked on the database?"

Sure. It's gone actually more smoothly than I thought it would. We were always upset if not all of our students give us every piece of information we want. We have – what would you say – close to 90% response rates to most of the items and even among the graduates, we got 70% of our grads. You know, we were sad it wasn't 90%, and so we will keep trying to work on strategies to increase that. But I think that is our biggest threat – is having people fall off over time and that's one reason we really try to cultivate their enthusiasm about the database as they go through the training; we try to help them feel that participation is important in contributing to knowledge, and we hope as we get more publicans and results out of the database and we disseminate that to people, they'll feel like it's something worth their time to continue to contribute to. Occasionally we ask a question in a way that we wished we would've asked it better, so the next year we have to change it, and those kinds of things that are just part of survey research. I can't think of anything else right now.

Thank you. We do have one more question – we have a couple more questions -- from the Q&A box. The next question is from Irene Sandvold, "Can you clarify what one big database for all of our programs would look like?"

[laughter]

Well, what we were envisioning, I guess this is an incremental approach, would be even if one other program forms a similar database, then we could make a data use agreement with them and merge our data, at least for certain projects. So, I really didn't see it is one big database, we did put that in our pie-in-the-sky slide, so I can see where that comes from. Ultimately, just in the same way that there is a lot of big data now available in other areas that we never thought would be aggregated in the way that it is. So I can't foresee how to actually do that, but I can foresee the next steps; and the next steps would be for more of us to collect this kind of data and of course doing that in a way that students are giving us permission to use it for research in the future and that is the incremental – crucial incremental step – so that when we do find a way to merge, to do the big merge down the road, we have something to merge [laughter]. In the meantime, I think it will be smaller collaborations.

Okay, we have a couple more I'm just trying to access this. I can't seem, Nolan, to open up the questions now. They were opening up before for me. Could you read the following questions?

If someone has a question that has not been answered that's in the Q&A, if you'd like to go ahead and follow the prompts provided by the operator, you could you ask them directly.

We do have one question in queue. Are you ready for that question?

Yes.

Alyssa Adams, your line is open.

Good afternoon. I just want to thank you for the thoughtful discussion and I wanted to mention that you did send out a link to the article in the announcement for the whole webinar where your bios are.

Great, that's good to know.

There is one more question from Pat Matthews-Juarez. The question is, "How long did it take you to create the database?"

We thought you would ask that, or someone would ask that [laughter]. I don't know the questioner. You want to take it? Our previous research analyst built the database over a month or two, I think, but she was doing other projects. I think she probably could have built it, if she spent full-time, she could have done it in two or three weeks. This is a person who is used to managing data, but it did not take an inordinate amount of time to build it. We had to write the consent forms and go through our first IRB process. If I added all that up, I would say it was about 40 hours of my labor. Brandi and I were talking just before the call about what to say when you all asked this, and we think, she is full-time here and probably spends less than 5% of her time maintaining this database. So that means adding all the many pieces of data that come in every year, really probably 100 pieces of data on each student, but it's all set-up so she can merge them in pretty quickly, so it's probably less than .05 of an FTE for the data manager. I hope that helps. That's about as good as guess as I think we can make.

Do we have any other questions?

Please use the operator instructions to ask your questions, that would be the most helpful way to do it at this point, rather than the Q&A box.

As a reminder, if you would like to ask a question, please press star one and record your name at the prompt. One moment while we wait for any incoming questions.

I'm currently showing no questions in queue.

I am showing that Maria P. has a question, but I can't see it. Maria P., could you press star one, please?

One moment. We do have the new question. The question is from Maria. Your line is open.

Thank you so much, operator. So, I actually am here with Irene and we have a couple of questions if that is okay. If we can take advantage of the airtime. The first question is what type of pre-graduation activities have you found that have been predictive of the post-graduate leadership positions? And, if it's okay if I just ask another question, the second question is do you also have high response rate to the SES questions that are a little bit more sensitive, like the clinical income as well as the sanction question. Those are the two easy questions. The other question is from Irene Sandvold: In terms of when you mention that you sometimes the response rate can wiggle around 70%, how do you follow-up and try to increase it back to your level of 90%? Can you provide more specific examples about the incentives, the amount of raffles or prizes etc., for the follow-up? Thank you so much.

Sure, I'll take the first question, which was the answer to a study that I posted during the presentation. That was actually a hypothetical study – we have not done it yet. It is the kind of thing we can do with this database. We only have -- we just collected our second year of data on graduates, so now – since we have about 90 students a year – we now have a large enough group that we can begin asking questions. So, we haven't actually done all these projects that I said we could do; we are just setting ourselves up to do them. So, I don't know the answer to that interesting question.

We would love to know the answer once you have it [laughter].

Okay, the next project on our list, right? As for the question about the SES, there is a lower response rate to that question, but out of our combined total of 174 responses, we had 150 students answer that question, so it is still relatively high. I think over the course of their time here we built a relationship trust with them and then also we have a really great alumni program that keeps up with them that helps maintain some of that trust, so that's good. And then as far as like reaching out and trying to keep the response rate high, the last couple years we have done some pretty big raffle prizes that we may not be able to keep up, but we have been raffling off branded apparel. Like apparel with the programs logo on it as a raffle prize, and so that's been very popular.

Yeah, and I think it wasn't so much that we saw them – well, I guess the entering students we get 95% probably on the first surveys and it stays close to that high throughout the program. When it dropped off to about 80% of the new grads. Where the 70% came from, I think, was that's how many – that's the proportion of the class that told us what they were making when they graduated. So, as you suggested, the hardest question to get people to answer and we got 70% to answer that. But, more like 80% answered our other questions, the easier questions such as, what type of practice are you in, what practice specialty, which are really important for outcomes the kinds of things that we care about, that HRSA cares about.

That's wonderful.

But a lot of it -- we do really talk it up. If our department chair or our Dean enters the building I try to catch him before he gets to the classroom and say, "Remember to mention the education research database and how important it is!" [laughter] We try to get a lot of people to say that they think it's great and it will contribute to making education better and making better clinicians. We say that every chance we get.

Thank you both so much for joining us today and presenting on this very important topic. We are at 3:01. I think that there might have been a couple other questions and what I'd urge you to do is to go ahead and please ask those questions to your GPO and we will get those and we can always return those responses to you. If there is anything urgent that you'd like to get some responses to as well, we will forward around that research paper that Dr. Morgan referenced, for everyone. Alyson, any closing remarks before we end the webinar today?

It looks like she doesn't have anything.

Sorry, guys, I'm talking on mute. I apologize. Thank you all very much for participating in this webinar. I think that this is a very timely topic. I hope you found some information that Dr. Morgan and Dr. Brandi – Dr. Leach also shared, helpful in terms of creating your evaluations, tweaking your evaluations that are currently being implemented. Like we said, once we get our website up and running with our PCTE toolkit we will be notifying you, and you will have access

not only to these webinar sessions, but also to other evaluation modules that will have all sorts of tips and tools and resources, bibliographies, things of that nature that you might find useful and we look forward to having you at our next webinar in two months. Take care, have a great day, and again if you have questions, feel free to reach out to me directly. Thank you, guys.

Thank you, this does conclude today's conference. All parties may disconnect.

>> [Event concluded]