David Murray:
Hello, my name is David Murray. I'm the NIH Associate Director for Prevention and Director of the Office for Disease Prevention. I want to welcome you to part five in our pragmatic and group randomized trials in public health and medicine course. This is a free, seven-part, self-paced online course presented by NIH. We provide the slides for each module, a complete set of readings for the course and guided activities for each module. This particular day we're going to talk about examples of group randomized trials. The course is designed for faculty, post doctorate fellows and graduate students interested in learning more about the design and analysis of group randomized trials. It's also designed for programs directors, program officers and scientific review officers who need to know more about these designs. Participants should be familiar with the design and the analysis of individual randomized trials and with concepts of internal and statistical validity, their threats and their defenses. We'd also like participants to be familiar with linear regression analysis of variants and covariates and logistic regression.

At the end of the course, participants will be able to talk about the distinguishing features of group randomized trials, individually randomized group treatment trials, and how they differ from individually randomized trials. Participants will be able to discuss the appropriate uses of these study designs in public health and medicine and for group randomized and individually randomized group treatment trials, to discuss the major threats to internal validity, to statistical validity, will be able to talk about the strengths and weaknesses of design alternatives and analytical alternatives and will be able to perform sample size calculations at least for a simple group randomized trial. Participants will also be able to discuss the advantages and disadvantages of alternatives to group randomized trials for the evaluation for multilevel interventions.

The outline of the course is shown here and today we will focus on some examples of group-randomized trials. The examples that I'm going to share with you today are from the health care assistants collaboratory. This is a project funded by NIH as a common fund project. These are funded as cooperative agreements. They're funded by a variety of different institutes and centers across campus, so they're not all from cancer or from heart, lung, and blood. They're really spread out. They're nine of them that are currently funded and as it turns out, eight are group-randomized trials. This is not terribly surprising to me because pragmatic trials are often done out in the real world working with existing systems and so very often employ group randomized trials designs. The eight that use the group randomized trial design are listed here in terms of the outcome that they're focused on. So one is focused on hospital acquired infections, another on colorectal cancer screening and so forth. I'm going to talk about several of these but not all of them and the first one that I'll talk about is based on colorectal cancer screening and it's called the Stop Colorectal Cancer Project.

This is a project that's led by Gloria Coronado and Bill Vollmer at Kaiser Permanente. The primary objective of the Stop CRC trial is to test the effectiveness of an automated, electronic medical record driven strategy to raise colorectal screening rates in safety net clinics. The

primary outcome for this trial is the proportion of targeted patients who complete fit kits during the first year of the intervention. This is a group-randomized trial. There are 26 federally qualified health clinics that have been randomized to intervention or control. They are affiliated with eight larger administrative networks and it's clinics that are randomized. They are randomized within networks. So we think of this as a stratified random assignment. Clinics within networks are randomized to conditions. We use the electronic medical record to drive the system level intervention and the electronic medical record is also the source of the data for this trial. Control clinics are going to roll out the intervention in the second year and consent was waived for participants because this was considered a minimal risk study.

This particular example illustrates a priori stratification in a group-randomized trial with clinic as the unit of the disbarment and it also illustrates a delayed treatment control condition. The analysis approach that's planned for this study is a weighted logistic regression accounting for clustering at the clinic level, adjusting for selected for individual and clinic level covariates. Individual level data will be weight by the inverse of the clinic size so that the resulting clinic means all have equal weight. This is consistent with the primary focus on clinic level outcomes. There's a paper published by Gloria Coronado in 2014 that provides details on the design and on the analytic plan. This particular analysis illustrates a mixed model and COVA approach, adapted for a dichotomous outcome variable.

Challenges that the Stop CRC Project has faced as it's gotten underway: As it turned out, there was going to be overlap of year one measurements in year two intervention roll out for control clinics so that had to be addresses. The use of real time electronic medical record tools wasn't always concordant with static randomization tables. There were implementation delays with the portable care act roll out. So all these challenges threatened the validity of the primary analysis. In terms of solving these real world problems, which happen in many, many group randomized trials; the team delayed the roll out of the intervention for the control clinics in the second year to address the overlap problem. They formulated a variety of sensitivity analysis to try to overcome lags in start up and so get a more accurate estimate of the true intervention affect, and they adopted the stepped wedge framework in which data from both years one and two, as well as a year prior to randomization is used to estimate separate start up effects for year one and the study stated effects in year two. So these are additional supplemental analysis, secondary analysis that will help with the interpretation. Adaptations are, as I said, often required over the course of a group-randomized trial because these studies are typically done in the real world.

The next example that I want to talk a little bit about is the chronic pain management study or PAC [spelled phonetically] and the primary outcome here is chronic pain management. This is project headed up by Linda Bar [spelled phonetically]. The biostatistician is again Bill Vollmer. This is another project from Kaiser. The primary objective is to test whether an integrated pain management program that's embedded within primary care will be effective in reducing pain, reducing opioid

use, improving or reducing health care utilization, and improves function for patients who have a condition that involves chronic pain.

This is a very timely project given all the interest these days in opioid use and so it's of interest to an awful lot of people. The primary outcome is the trajectory of change in self reported pain scores over the first six months of the intervention. So this is a different kind of outcome that's a slope that's being estimated using data and then we estimate a slope for each person and compare those across groups and across conditions to see if there's an intervention effect. This is an example of a stratified group randomized trial. The strata are the regions of the Kaiser Permanente health plan. Physicians are the unit of randomization in this study, not clinics but physicians themselves. The electronic medical record is used to screen and identify potentially eligible patients. The patient -- potential patient list are then vetted with their primary care providers and verbal consent is obtained from patient prior to randomization.

This is another study that illustrates the stratified group randomized trial and in this case we have the physician as the unit of assignment. In the first example we had clinic as the unit of assignment. The analysis plan is a two-stage approach. The first stage would calculate a slope in individual pain scores in each patient. Then those slopes would be analyzed using a mixed model and COVA, adjusting for selected individual and cluster level variables including baseline pain score and there's a paper from Linda Bar published in 2012 that provides details on the rationale for this approach. This analysis plan illustrates a two-stage analysis, also reflects regression adjustment for covariates, an approach that's very common with group-randomized trials.

Challenges that the PAC study faced part of it was weaving a very complicated, multimodal intervention into the fabric of the usual care in the intervention -- for the intervention providers. Everyone was doing lots of different things that they hadn't done before, so this study involved redeploying or hiring new clinical staff for intervention roles that weren't necessarily well aligned with the existing health plans structure or the traditional activities of the office. This study expanded the use of the electronic health record. There was a challenge creating a scalable training model. Costs are often an issue. IRBs are often also challenging. We think that that situations going to get easier going forward because pretty soon multisite trials are going to require -- at least funded by NIH are going to require a single IRB rather than multiple IRBs. These kinds of challenges just reflect the fact that pragmatic trials are not easily, especially if you're doing some thing for the first time in a new system with new methods and the PAC team had to work a lot of these issues out. They adapted the intervention structure to accommodate the clinical workflow. They had to redefine some clusters by grouping PCPs together due to smaller than expected numbers of consenting patients for some providers. They had to delay the startup in some regions until the systems were ready. They had to shift some of the sample size between regions to reflect what was possible. So there was a variety of adaptations that were required to adapt to the circumstances that they found.

As the next example I want to describe the TSOS [spelled phonetically] project. This involves management of post-traumatic stress disorder in trauma patients. This is a trial directed by Doug Zatzick and the biostatisticians is Pat Haggerty [spelled phonetically]. All of these folks are at the University of Washington. The primary objective is to explore the intervention effect in patients with pre-injury chronic medical conditions and the primary outcome is post-traumatic stress disorder symptoms. This study uses the stepped-wedge design. Now we haven't' talked in much detail about stepped-wedge design so I'll pause and give you a little more information on this one. This kind of design. A stepped-wedge design has a group -- a series of groups that are going to receive the intervention. Groups are selected in random order to receive the intervention following a baseline period. So all of the groups provide baseline data. All of the groups provide intervention data over the course of the study. In a stepped-wedged fashion, the interview is introduced into all of the groups. In this case there are 24 Level One trauma centers that are randomized to four waves. So in the first wave -- at the beginning everybody provides up control data and then the first six trauma centers receive the intervention. Awhile later another six and so forth. So it's a stepped-wedge design.

The goal was to recruit 960 patients with PTSD, 40 patients per trauma center. All of the trauma centers recruit both control and intervention patients by virtue of the stepped-wedge design. All of them begin by recruiting controls and the data are collect at baseline three months, six months and 12 months. The intervention is turned on at each trauma center based on the stepped-wedge design. One of the advantages of the stepped-wedge is that all of the trauma centers get trained on doing the intervention over the course of the study but the design does add analytic complexity.

So one of the features of this project is that it does illustrate the stepped-wedge design. The intervention and control comparisons will be made for PTSD symptoms. That's the primary outcome but the time will also look at alcohol use and depression. Subgroup analysis will focus on participants who have traumatic brain injuries or pre-injury medical conditions. There's a paper by Hughes [spelled phonetically] published in 2015 that provides a discussion of the analytic issues in the stepped-wedge design and illustrating mixed effect regression approaches with adjust for covariates and here's the reference for that paper.

Challenges: there as considerable variability among the sites. They varied in the extent of violent injury and of course there's more PTSD with violence. The sites varied on a variety of other characteristics. The stepped-wedge design actually helps with this variability. Implementation challenges: The American College of Surgeons mandated activities for PTSD screening and intervention, so all the sites do want training. The solution -- as I said, the stepped-wedge design actually addresses the site variability by the providing the intervention in every site so that every site contributes both control and intervention observations. The team dealt with the requirement from the professional society by providing the training to all of the sites at least by the end of the study.

Another example that I want to talk about is a study called Pieces [spelled phonetically]. This is a project that focuses on management of chronic conditions -- multiple chronic conditions. It's run by Michael -- Miguel Vazquez and Chul Ahn is the primary statistician out of the University of Texas Southwestern Medical Center. The primary objective of this study is to evaluate the management of patients with chronic kidney disease, diabetes and hypertension with a clinician support model that's enhanced by technology support and that's the Pieces intervention compared to standard of care. The primary outcome is one year all cause hospitalization among participants that are in the trial. This is an example of a stratified group randomized trial. They have four health care systems what almost 250 clinics and more than 35,000 patients who were available for the study. Within each healthcare system, clinics or practices sites will be randomized either Pieces or Standard care. And every patient assigned to a clinic or particular practice site will get the treatment based on the randomization of the clinic or practice site. As I said, this illustrates the stratified group randomized trial design. IN this case, the clinic or practice site is the unit of assignment.

The primary analysis in Pieces will be the generalized Mantel-Haenszel testing procedure, published in paper by Allan Donert in 1992. This procedure will be applied to detect any different in hospitalization rates between Pieces and the control condition. Secondary analyses will use model-based methods. Mixed model logistic regression to assess intervention effects on hospitalization rate, controlling for clustering in-patient and other covariates. Cox models to assess the intervention effect on time to hospitalization with the shared frailty to control for clustering.

This study illustrates the use of a non-parametric approach to the primary analysis and model-based approaches to secondary analysis. Challenges faced by the Pieces project are summarized here. There was a challenge getting consent waivers but that was worked out. There was a challenge with the heavy workloads among the participating centers. Streamlining the clinical workflow at each of the sites. Certainly there were competing priorities for a variety of things but especially with some of the technology development that was required for this project. There was a slow approval process at one of the study health care systems that slowed things. And then there was training of all those PCPs and staff at each of the sites. These kinds of logistical problems are common in pragmatic trials, especially done in healthcare settings.

The team is addressing those and so we'll move to a summary of these examples. Group randomized trials, individually randomized group treatment trials can be applied in a wide variety of settings for a wide variety of primary outcomes. Group randomized trials should be avoided if individual randomization is possible with no threat of contamination or interaction among participants post-randomization. But if you've got the threat of contamination or there's going to be interaction among participants post-randomization, the group randomized trial or individually randomized group treatment trials provide the strongest design that you can have. These studies are often conducted in settings where investigators have limited control and teams should included

experts in the settings and operations so that they can address those issues as they arise, not just in the interventions and outcomes.

I want to thank you for participating today in part five of our course in pragmatic and group randomized trials in public health and medicine. I want to draw your attention to our website where you can provide feedback on today's material. You can download the slides that you've seen on this module, the references for the entire course and suggest activities to follow this particular segment. You can view this module again, or you can do the next module in the series, part six, a review of recent practices. If you have any questions, please send them to grt@mail.nih.gov. And thank you very much for your attention.

[end of transcript]