Jill Harrison, PhD:

Hi, this is Jill Harrison, Executive Director of the National Institute on Aging IMPACT Collaboratory at Brown University. Welcome to the IMPACT Collaboratory Grand Rounds Podcast. We're here to give you some extra time with our speakers and ask them the interesting questions that you want to hear most. If you haven't already, we hope you'll watch the full Grand Rounds webinar recording to learn more. All of the companion Grand Rounds content can be found at impactcollaboratory.org. Thanks for joining.

Vince Mor, PhD:

Dr. Li, I'm delighted to have occasion to speak with you about your wonderful Grand Rounds that you did yesterday. We had lots of engagement of the audience, and I have several questions that come up, which I think it would be wonderful if you could answer. Your work on basically design issues in the execution of cluster randomized trials are absolutely pertinent to the work that the IMPACT Collaboratory has been doing over all these years, and as a member of the design and statistics core, that would make lots of sense for you to have been involved in those. The first question that emerges is from Tom Travison, your colleague, who's also a member of the design and statistics core.

He says, "Thank you for your great work, your great Grand Rounds." He asks in question one, "In a situation where the overall effect and the presence or absence of meaningful heterogeneity of the treatment effect, do you recommend first testing for heterogeneity of treatment effect and then simplifying to an overall effect of the heterogeneous treatment effect and see whether that's judged to be shown by a significance or absence, or do you test that as part of an overall interaction term?”

Fan Li, PhD:

Thank you, Dr. Mor, for the question, and it's a pleasure to be here just to share my perspectives on these questions. Just to start, I think for Tom's question, in general, I think as I'm dealing with study design issues here, we generally recommend the pre-specification of all the testing that's going to be conducted in a specific pragmatic trial. There is a consequence that's not so favorable in doing a step-down test in a sequential order by testing for significance first and then decide what you're going to do next after seeing that evidence.

I think you are risking a type one error rate issue, and so we generally do not recommend that approach and would much prefer to pre-specify the analysis that's going to be carried out in a specific study.

Vince Mor, PhD:

Under that circumstance, what you're essentially saying is that you'll be testing for the main effect, and then you would introduce an interaction effect to test for the heterogeneous treatment effect simultaneously in your primary analysis. Is that correct?

Fan Li, PhD:

Not exactly, because I think sometimes the primary analysis would be decided based on the primary objective of the study of interest. I think what I would recommend is that if indeed both effects are of interest, then we would be doing both simultaneously, and then the consideration of whether we need some sort of multiplicity adjustment would factor into the consideration. In many other cases where one of these two could be the primary pursuit of a specific study, we can still use the same model, but we would have a priority in testing either the overall or heterogeneous effective treatment.

Vince Mor, PhD:

I see. In other words, if your primary hypothesis is based on the overall, which is where your sample size was actually generated in terms of the number of clusters and the number of individuals per cluster, you would test that first, and then you would test the heterogeneous effect with some kind of adjusted Bonferroni-like correction.

Fan Li, PhD:

Right. Sometimes that's going to be the secondary analysis that's complimentary to the overall finding that we want to see in a primary analysis. And even in that situation, the discussion in my presentation would offer a context under which we could interpret the secondary analysis. We can see, given these study resources we have to power the overall analysis, do we actually have enough power for those secondary supplementary analysis for testing interaction? If not, it's also a way to make sure we have the correct interpretation of that analysis and not to over-interpret the findings therein.

Vince Mor, PhD:

Yeah, that would be completely ad hoc or a supplemental analysis done, which is exploratory in that sense.

Fan Li, PhD:

I mean, in some sense, you could specify those information or power information even in the beginning of the study, but you would caution that unlike the primary analysis, these additional analyses may not be sufficiently powered. We do interpret them with caution.

Vince Mor, PhD:

Okay, all right, great. Related to that, you started your talk yesterday with some discussion about this kind of heterogeneous treatment effect idea would be particularly important sometimes in doing work related to health equity to see whether or not the introduction of these interventions would be equally efficacious or alternately efficacious in a minority population, let's say, of African American study subjects drawn from the overall study effect.

Since I've done work like this in the nursing home context and there are lots of data from lots of literature suggesting that not just in nursing homes but also in ambulatory care settings, in primary care practices, and even hospitals, the distribution of minority or White versus Black patients is not random. There will be some distribution of those individuals who are of interest to a heterogeneous treatment effect question that will not be distributed randomly over all of your clusters.

Particularly in nursing homes, but also in primary care practices, there will be some practices or clusters where there'll be no Black patients, some where there will be, perhaps, all Black patients, a small minority, and the rest will be some unequal distribution of individuals across the remaining clusters. Could you comment on how that circumstance affects the choice of looking for a heterogeneous treatment effect in the first place, using it as an individual level covariate or using it as a cluster level covariate? Because from my vantage point, this gets very complicated very quickly.

Fan Li, PhD:

Thanks, Dr. Mor. I think it is a very important consideration. I think, almost in every study that sometimes you do have a substantial amount of, let's say, between cluster differences in how those attributes get distributed. I mean, this is raised as an example, but I think other characteristics we are

seeing maybe similar patterns again and again in different studies. I think this is particularly relevant to the concept of intracluster correlation coefficient of the covariate. I think I should be a little bit more clear on what that means. That quantity actually measures the extent to which the distribution of the covariate or the variable that we are interested in measuring heterogeneity effect distributed across the clusters.

For instance, if we have a very extreme situation where we either have clusters where there are going to be White population completely or other clusters with non-White population completely. This is an extreme separation, in which case, the intracluster correlation coefficient of the race variable becomes one, which is its upper bound or the largest amount possible value that intracluster correlation can possibly take. On the other end of the spectrum, we could imagine a scenario, even though highly unrealistic, where the race subvariable is randomly distributed across the clusters.

In that specific situation, the intracluster correlation coefficient of this variable becomes zero. Whatever happens in between these two extremes signals an intracluster correlation coefficient of the race variable in between zero and one. And if it moves close to one, we have a lot of between cluster heterogeneity or differences in the distribution of the race variable. In the approach I have introduced yesterday, we do have a key element in the sample size expression that controls for the amount of heterogeneity or between cluster differences on this information that we are interested in.

We would be able to accommodate that. In real practice as to how this value, what this value is in specific application, I think that's a case by case discussion. We would be able to inform that if we exploit existing databases. For instance, I think we have similar resources in the IMPACT Collaboratory that we are tapping into recently.

Vince Mor, PhD:

Great, great. Thank you very much. On a related issue, but somewhat different, effect heterogeneity may be explained post hoc by differential clustered level performance in implementing the intervention. This is a facility level factor, but these are almost always post hoc kinds of analyses. Do you have any suggestion for those who are trying to anticipate that their participants, their clusters will not be comparably implementing the intervention?

How can you take that into consideration in the design of a research project? Imagine that some proportion of your folks are just not going to do it very well, some will do it greatly, and some will be the great unwashed in the middle. How do you anticipate something like that in the design?

Fan Li, PhD:

Yeah, that's a very important question. I think we are seeing a lot of those in the pragmatic trial setting where differential providers or clusters do have a different time schedule or a different practice in introducing the intervention, which ends up being very different between or across clusters. But on the other hand, we need to realize that the provider level implementation strategy or practice is a little different from the covariate or effect moderation I was focusing on yesterday.

I was focusing on the baseline information, which was taken before the randomization was even performed. In your case, which is practical and highly important, the complication is that the provider level implementation is measured or assessed after the intervention was introduced. This does create a complication where we are talking about potential effect heterogeneity that could be possibly explained by a post randomization variable that could be affected by whether the provider receives the intervention or not.

I think in this particular situation, it has some resemblance to the complications arising from non-compliance or non-adherence. In which case, it's not only about effect moderation or modification, it's more about how do we actually measure the actual intervention effect to get its efficacy evidence versus the effectiveness evidence, which we could measure just by getting the randomization assignment information in the particular study, ignoring the implementation details downstream. I think there's a potential to integrate what I have presented with the provider level implementation challenges.

However, the challenge statistically would be we have to introduce a framework to define the quantity or summary measure of interest that could take into account the differential implementation at a provider level. I think that would be the first step to make sure we agree on what we can call an intervention effect of interest, after which we could explore down the road whether that intervention effect could be explained by some other baseline covariates of interest. I do think that these are two related, but somewhat different complexities, and there is a future point where we could combine these considerations to address complexities in pragmatic trials.

Vince Mor, PhD:

Great. Thank you. That's actually very well explained. It's the pre versus during, what's post random assignment, and that becomes a different kind of confound for post hoc exploration. Thank you very much. Dr. Li, thank you very much for a wonderful Grand Rounds yesterday, and this has been a lot of fun. I'm hoping actually, for the second question related to race and ethnicity, we could actually do some work together on this. It would be very exciting. Thank you very much for your time. Very much appreciate it.

Fan Li, PhD:

Thank you so much, Dr. Mor. It's my pleasure to be here and to discuss these questions with you. Thank you.

Jill Harrison, PhD:

Thank you for listening to today's IMPACT Collaboratory Grand Rounds Podcast. Please be on the lookout for our next Grand Rounds and podcast next month.