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.

Hello everyone. Welcome to today's podcast. I'm Jill Harrison, one of the executive directors for the IMPACT Collaboratory. On behalf of IMPACT leaders, Drs. Susan Mitchell, Vince Mor, and Ellen McCarthy, thank you so much for joining us today.

I'm joined for this podcast by Dr. Eleanor Murray. She's an Assistant Professor of Epidemiology at Boston University School of Public Health. Dr. Murray, thanks so much for joining us for the podcast.

Eleanor Murray, PhD:

Glad to be here. Thank you for inviting me.

Jill Harrison, PhD:

Yesterday you conducted a Grand Rounds on the topic of causal inference in pragmatic trials. Can you just remind our listeners, what is causal inference in pragmatic trials?

Eleanor Murray, PhD:

Yeah. So causal inference is really the umbrella term for the idea of trying to estimate how much an exposure or a treatment, how much of an outcome it causes or prevents. So it's really about estimating the amount of change we'd see if people got one treatment versus another treatment or one treatment exposure versus another exposure. Pragmatic trials are a great tool for doing this if we're careful about how we design them. The randomization piece helps us avoid confounding for treatment assignment.

And then a lot of features of pragmatic trials are designed around trying to improve the relevance of the answer to a wider population of individuals with a condition, compared to, say, an explanatory randomized trial, which might be a very, very specific subset of individuals enrolled. In thinking about pragmatic trials from a causal inference perspective, one of the key pieces there is to think about what actually is the decision that you want to inform with your pragmatic trial, and are you actually asking a question that answers or provides some evidence for that decision with your design and analysis plan?

Jill Harrison, PhD:

Thank you so much for that summary. Yesterday in the Grand Rounds, there was a lot of interest in this topic and several chat questions that we didn't get to, so I'd like to pose them to you now. These were from our Grand Rounds listeners.

"Do you have any advice on addressing immortal time bias when the exposures do not begin at the same time? For example, after knee replacement, patients use pain medications immediately, but may not use physical therapy for a couple of weeks."

Eleanor Murray, PhD:

Yeah, this is a great question because immortal time bias is a problem that crops up a lot when we're trying to analyze observational or secondary data, but in pragmatic trials, this should be a little bit easier to handle. That's because we can start the clock of when we're looking at the outcome, we can start it right at randomization. So long as people are entering our trial at the same time point in their illness, or

for example, entering at the time of knee replacement or at a specific time after knee replacement, so long as that's the same for people in both or all of our trial arms, then we don't expect to have immortal time bias in a pragmatic trial.

If you think about it, in the trial, we count outcomes that occur based on which arm they happen in, regardless of whether or not somebody has, for example, started physical therapy yet. When we're thinking about an observational study or electronic medical records, for example, like existing data, it can be a lot harder to know when to start that follow-up period. So what we often have happen is that patients who have some serious complication from their surgery never get to the point of starting physical therapy, and then that counts them as someone providing evidence to "not physical therapy". But in a trial, that person could have provided evidence in the physical therapy arm of the trial, it's just that they never actually got there.

So when we're concerned about this immortal time bias issue, it really comes about from, are there people who might experience something that's part of the exposure or a competing event like dying or becoming ineligible for treatment before treatment happens, and are you doing anything different about those people between your two trial arms? So long as those people are included with the same likelihood in all your trial arms and you're not systematically excluding a group from one or the other trial arm, then you should be good with immortal time bias.

Jill Harrison, PhD:

Thank you so much. On a related question, can you talk a little bit more about why we should bother with per-protocol analysis and not just focus on the intent to treat?

Eleanor Murray, PhD:

Yeah. So this is another place where there's often a lot of confusion around pragmatic trials, because in a lot of ways, the goal behind pragmatic trials is to try to target a much more clinically relevant question. For example, how well does a new treatment perform versus standard of care? The intention to treat effect really only tells us for sure about what the effect of randomization is under the particular pattern of non-adherence or adherence that we see in our trial.

It's often the case that the levels of adherence in a trial are different than they will be after the trial is completed. For example, if we're trying in our trial something that people are a little bit unsure about, they may adhere less, but once our trial's completed, if it provided strong evidence that the treatment works, people may be more likely to adhere. That intention to treat analysis can actually end up providing a slightly skewed view of what the actual benefit of treatment is.

This is particularly important when we're thinking about harms, because even if our comparison is a placebo, which is typically not the case in pragmatic trials, but sometimes people may just have no treatment as the comparator. In that case, the intention of treat underestimates the real effect of the treatment when there's non-adherence. If what we're concerned about looking at is a safety or adverse outcome, that's not really something we want to be underestimating.

Even more when the comparison is in separate active treatment, for example, usual care, non-adherence can actually end up making the intention to treat either bigger or smaller than the true effective treatment, and we don't necessarily know which one. This is why we recommend people think about the per-protocol effect, but it's important to understand that when we talk about the per-protocol effect, we're not talking just about restricting the analysis to people who adhere and then looking at the same type of analysis that we would for the intention to treat, but just in that subpopulation. If we do that, we can end up with a really out-there answer, which is strongly affected by

the fact that people who adhere to their treatment are going to be different in many ways from people who don't adhere to their treatment.

So when we think about estimating the per-protocol effect, what we want to do is use this effect estimation causal inference framework to think about how do we estimate what the effect would've been in the trial if everyone had adhered. This basically uses information that we have measured and know about people who did and didn't adhere and compensates for what happened with people who didn't adhere by trying to find similar people among the adherent people and up weighting them to correct. It's a little bit more technical than I got in the talk, but that'll give some extra information to people who listen to the podcast as well.

Jill Harrison, PhD:

Fantastic. We always love extra knowledge nuggets for our podcast listeners. It's very much appreciated. One additional question from a listener yesterday, "I'm still trying to understand how do you test that loss to follow-up as random or not? A simple approach would be to compare across different confounders if they are balanced or not within loss to follow-up, treated, and control groups, but concluding that there is some unbalance, it could be due to loss to follow-up or to adherence or to something else, right?" Can you help clarify?

Eleanor Murray, PhD:

Yeah, so this is the other big challenge. In pragmatic trials, we have the issue of adherence, and we also have this issue of loss to follow-up. I think it's important when we talk about loss to follow-up to acknowledge that when people stop coming in to trial visits or move away, drop out, withdraw from a trial, this can cause problems even for our intention to treat analysis. For example, if you imagine the extreme case, if everybody who was taking treatment decided they didn't want to participate in our trial anymore, we'd have no information on what happens in the treatment arm and we'd end up really not knowing anything in our intention to treat comparison. So this is not something that's only specific to pragmatic trials or only specific to the per-protocol effect. It's really a bigger randomized trial and observational study issue.

So here, loss to follow-up is not going to be that much of a problem if it basically happens randomly between the different types of people in our study. The first question we need to ask ourself is, how does loss to follow-up differ between the trial arms? If we see that people in the usual care arm are dropping out much more than people in the treatment arm, that's a big red flag. That means that we're going to need to do something to correct for loss to follow-up because if it's happening more commonly in one arm than the other, then that very likely means that one type of person is more likely to drop out than another. Because, basically, randomization tells us that the types of people in both arms of our trial should be roughly the same.

When it comes to getting more specific in terms of what actually should we be controlling for, what types of confounders of loss to follow-up do we need to look at, there's a little bit of a tension between the fact that we can't necessarily, as the question askers say, we can't necessarily know for sure that an unbalanced is coming from non-random loss to follow-up, or whether it's just a statistical difference that doesn't actually have an important meaning. The best approach is to really think about, first, does loss to follow-up differ between your trial arms. Then is there a set of variables that you think might be more common among those people who are versus aren't lost to follow-up, with that second set, which you can base partly on your data, but you also want to kind of base on what makes sense in terms of what you know about people in your trial. Then think about the subset of that, that is associated with your outcome, and those are the variables that are really going to be important to control for.

It's also worth noting that a lot of the ways to adjust for adherence can also be used to adjust for loss to follow-up, and you can do both types of adjustments together. Sometimes you may find that, for example, maybe older individuals are less likely to adhere and more likely to be lost to follow-up. Well, we can address both those problems so long as there are at least some older people who stay in our trial and take their medication. It's not necessarily hugely important to say, well, it must be due to this or it must be due to that. We just need to know who are we more likely to have an outcome measurement for and use that information to get the best answer we can.

Jill Harrison, PhD:

Thank you so much. The mission of the IMPACT Collaboratory is to grow the field of pragmatic trials, particularly for dementia care. So as our last question, for folks entering the field of pragmatic trials in dementia care or traditional trialists entering the realm of pragmatic trials, what resources or training materials would you point them to if they want to grow their skills in not only designing trials, but specifically in causal inference in pragmatic trials?

Eleanor Murray, PhD:

Yeah. So there's a nice paper on the per-protocol effect in pragmatic trials in the New England Journal of Medicine from a few years ago, written by Miguel Hernan and James Robins. I would definitely point readers to that, which gives a nice summary. Those authors, Miguel Hernan and Jamie Robins, also have a textbook, which is available online, but also coming out in print soon, which is another great resource, and I would absolutely point to that.

For readers who want to delve a little bit more off the beaten path, we do have some preliminary guidelines for pragmatic trials and causal inference, which are available on Archive, but they have not yet managed to make it through the stack of to-do things out through the publication process. So those might be something people would want to think a little bit carefully about and see whether or not they agree with.

Jill Harrison, PhD:

Well, Dr. Eleanor Murray, Professor of Epidemiology at Boston University School of Public Health, thank you so much for your time today. I want to remind our listeners that they can also follow you on Twitter at @EpiEllie. Again, just extend our gratitude for your time, not only in the Grand Rounds, but in today's podcast. Thank you so much.

Eleanor Murray, PhD:

Thank you for having me.

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.