Journal of Physiotherapy 67 (2021) 147–149

Journal of PHYSIOTHERAPY

journal homepage: www.elsevier.com/locate/jphys

Appraisal

Research Note: Estimating the complier average causal effect when participants in randomised trials depart from allocated treatment

Randomised trials are frequently conducted to test the impact of interventions on outcomes in physiotherapy. For example, patients recovering from surgery may be randomised to receive a standard or a novel rehabilitation program, or knee osteoarthritis patients may be assigned to wear a certain type of shoe.[1,2] Consider the Telecare trial: participants with knee osteoarthritis were randomised to an existing telephone support service (the control arm), or five to ten telephone consultations with a physiotherapist in addition to the support service (the intervention arm).[3] One of the primary outcomes was physical function at 6 months after randomisation. Due to randomisation, the groups were expected to be similar with respect to measured and unmeasured participant characteristics, so any difference in the outcome between groups could be interpreted as being due to randomisation to different treatments.[4]

But what if participants depart from their randomised treatment? In Telecare, some of the intervention arm participants had fewer than five telephone physiotherapy sessions. How then can the effect of actually receiving the full course of the intervention on outcomes be estimated, as opposed to the effect of being randomised to receive the intervention? This research note discusses how the causal effect of full compliance with an assigned treatment can be estimated. Common (and frequently flawed) approaches, such as the 'per-protocol effect', are contrasted with an approach within the paradigm of causal inference. This follows on from a previous Research Note introducing the framework for causal inference.[4]

Causal diagrams for randomised experiments

Directed acyclic graphs are a type of causal diagram used to graphically display assumptions about causal relationships between variables.[5] Figure 1 displays a causal diagram for the Telecare trial: nodes represent variables, with directed edges (arrows) representing the presence and direction of causal relationships. The lack of an edge between two variables encodes the strong assumption that there was no causal relationship between those variables.

In Figure 1, the randomised group is a cause of 'Number of physiotherapy sessions': participants assigned to the intervention arm were expected to have higher numbers of telephone physiotherapy sessions than those assigned to the control arm. This causal diagram indicates that there may have been other variables – some measured, some unmeasured – that were causes of both the actual number of sessions and the outcome. For example, perhaps participants with lower levels of function at baseline participated less in the program, and were more likely to have worse functional outcomes at follow-up. Such variables are often referred to as 'confounders', although the definition of a confounder is broader than this.[6] There are likely to be several such variables that remain unmeasured: for example, there may have been certain unrecorded comorbid conditions that influenced both the number of physiotherapy sessions and overall function levels.

Intention-to-treat and alternative effects

In an analysis of a randomised trial, the primary research question will usually be in terms of the differences in the outcome between randomised groups, often referred to as the 'intention-to-treat' effect.[7] The intention-to-treat effect compares the outcomes of all participants that were randomised to the control arm with the outcomes of all participants who were randomised to the intervention arm, regardless of whether they later departed from their assigned treatment. The intention-to-treat effect is the only comparison for which the groups are guaranteed to be exchangeable, by virtue of randomisation. When compliance is not perfect, effects other than the intention-to-treat effect may also be of interest.[7]

For example, a 'per-protocol' effect may be estimated by excluding participants who did not receive their assigned treatment from the analysis.[8] However, by restricting the analysis in this way (that is, dropping those participants who did not receive their allocated treatment from the dataset) the benefits of randomisation are lost, due to the confounding of the relationship between the 'treatment actually received' variable and the outcome variable (as shown in Figure 1): as such, the per-protocol approach is not recommended. Other approaches have been considered, such as an as-treated analysis, where participants are re-categorised according to the treatment they actually received.[8] Such approaches will often compare groups of participants that are not comparable; the next section provides insight into why this is the case.

Causal effects when participants do not receive assigned treatment

Potential outcomes under each treatment

In a randomised trial where participants are randomised to the control or intervention arm, each participant has two *potential outcomes*: the outcome that would have been observed had they been assigned to the control, and that which would have been observed had they been assigned to the intervention.[4] Ideally, the two potential outcomes should be compared for each participant. In reality, only one of these potential outcomes is ever actually observed for each participant: that which corresponds to the treatment they actually received. These participant-level causal effects cannot be estimated, but provided that certain assumptions are valid,[a] *average* causal effects can be estimated: the differences in the average potential outcomes across all participants.[9]

Classifying participants by their potential compliance with each treatment

Consider the Telecare trial: the number of telephone physiotherapy sessions each participant had was recorded. A participant

https://doi.org/10.1016/j.jphys.2021.02.002
1836-9553/© 2021 Australian Physiotherapy Association. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

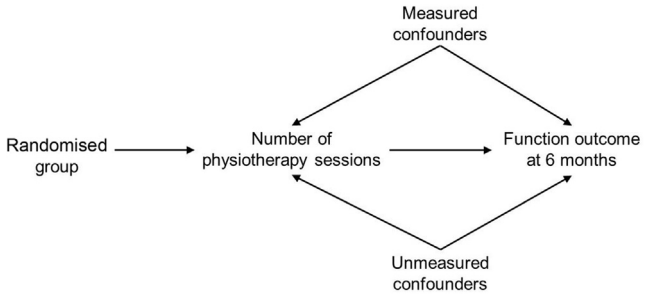

Figure 1. A causal diagram for the Telecare study. Participants were randomised to receive a standard telephone support service, or telephone physiotherapy sessions in addition to this telephone support service. The number of these sessions that each participant undertook was recorded, and the final self-reported function outcome was measured at 6 months after randomisation. Some confounders of the relationship between the number of physiotherapy sessions and the outcome at 6 months were measured (eg, baseline function levels, participant age); others may not have been measured (eg, socioeconomic status, comorbid conditions).

was considered to have received the full intervention if they had a minimum of five sessions. A participant is said to have complied with their assigned treatment if their randomised group matched their actual treatment. When participants may depart from their assigned treatment, we can consider potential *compliance* in addition to potential *outcomes*.

Each participant has two potential compliances: their compliance were they assigned to the control and their compliance were they assigned to the intervention. As was the case for potential outcomes, one of these potential compliance outcomes could be observed for each participant: that which corresponded to the group to which they were actually randomised. However, it is useful to consider both potential compliance outcomes for each participant (ie, their potential compliance under each of the two treatment allocations), and to classify or stratify participants according to these hypothetical levels of compliance. Table 1 shows this stratification: stratifying participants on the basis of hypothetical compliance under assignment to each arm. Participants would be 'always-takers' if, no matter which group they were assigned to, they took the intervention arm treatment; 'compliers' if they always complied with their assigned treatment; 'defiers' if they always took the opposite treatment to that which they were assigned; or 'never-takers' if they never took the intervention arm treatment. This stratification is referred to as a 'principal stratification', where the term principal denotes that this classification does not depend on the actual treatment allocation of a participant.[10] Since only one of the hypothetical compliance levels is known for each participant, which of these strata any individual participant may belong to usually cannot be identified: this classification is latent or unknown.

This principal stratification is used to show why the per-protocol effect – where only those participants who received their assigned treatment are included in the analysis – is a flawed estimate of any relevant, well-defined population quantity. The groups being compared in the per-protocol effect are those highlighted in blue in Table 1. Participants who were assigned to the intervention and received the intervention will be either always-takers or compliers; participants who were assigned to the control and received the control will be either compliers or never-takers. If the always-takers and the never-takers differ, the per-protocol effect compares the outcomes for groups of participants who are not, in fact, comparable. Thus, the per-protocol effect does not have a valid causal interpretation: any difference that is observed may be due to differences between always-takers and never-takers, rather than due to the intervention.[11]

Table 1
Principal strata for compliance in a randomised trial. Participants may be randomised to the intervention or control arm, and this table describes hypothetical treatment receipt under both possible assigned treatments. A participant is said to 'comply' with their assigned treatment if treatment assigned matches treatment received. Cells shaded in blue indicate when a participant receives the assigned treatment; unshaded cells indicate when a participant departs from the assigned treatment.

Participant classification	Assigned to intervention arm	Assigned to control arm
Always-takers	Receives intervention	Receives intervention
Compliers	Receives intervention	Receives control
Defiers	Receives control	Receives intervention
Never-takers	Receives control	Receives control

Complier average causal effects

If the per-protocol effect is not to be recommended, how can researchers estimate the effect of the treatment that participants actually receive on outcomes? The answer is to estimate an effect that has a causal interpretation by restricting the analysis to a single principal stratum.[10] Among those participants who were observed to have complied with their treatments (highlighted in blue in Table 1), there is a subset of participants who are comparable: the compliers. The complier average causal effect restricts attention to this principal stratum of participants: those participants who would always comply with the group they are randomised to. The complier average causal effect compares the outcomes between the compliers who happened to be randomised to the intervention arm and those compliers who happened to be randomised to the control arm. Since the principal stratum of a participant is a pre-randomisation characteristic of that participant (albeit an unknowable one), randomisation ensures that compliers assigned to each of the two arms are exchangeable (ie, these groups are expected to be similar). Thus, the complier average causal effect is a 'randomisation-respecting' estimate.[12]

Estimating the complier average causal effect

Even though it is unknown which participants are compliers, the complier average causal effect can be estimated using data from a randomised trial, provided certain assumptions are valid. The complier average causal effect can be estimated if the assumptions in footnote [a] are valid, if the randomised group can only change the outcome through the treatment actually received, and in addition there are no defiers.[13] The assumptions in footnote [a] will generally be valid for a randomised trial: randomisation ensures no confounding of the relationship between the assigned group and the outcome and between the assigned group and compliance; and all participants have a chance of being assigned to either treatment group. The final assumption in footnote [a] will be valid provided that the treatment has been clearly and precisely defined.

The assumption that a randomised group can only change the outcome through the treatment received is known as the 'exclusion restriction' assumption. This assumption requires that the only way that the intervention changes outcomes is through engagement with the intervention: that is, being randomised to either arm does not in itself lead to a change in the outcome. Participants must actually receive the intervention in order for the intervention to have an effect. For the Telecare trial, a simplifying assumption was made that five or more telephone physiotherapy sessions would be required for these sessions to have an impact on the outcome (so that receiving the intervention was equivalent to participating in five or more sessions). Participation in four or fewer sessions was assumed to have no effect on the outcome: whether this is a valid assumption requires subject-matter expertise. Methods that allow for degrees of compliance instead of such all-or-nothing compliance are available.[14,15] The assumption of no defiers is valid if participants assigned to the control arm can never access the intervention. For the Telecare trial, only participants assigned to the intervention arm had access to the telephone physiotherapy sessions, and thus this assumption is valid.

If all of these assumptions are valid, then a point estimate of the complier average causal effect is the intention-to-treat effect divided by the proportion of participants who are compliers.[14] If there are no defiers and participants in the control group cannot receive the intervention, then the proportion of compliers can be estimated as

the proportion of participants in the intervention arm who take the intervention. This approach is very simplistic. In practice, more complex approaches are used, including instrumental variable regression. Instrumental variable regression uses a two-stage least squares method to estimate the complier average causal effect and a valid 95% CI.[13,14] This approach fits two models: the first is a model for the treatment actually received, predicted by the randomised group and other relevant covariates. The treatment predicted by that model for each participant is obtained and the second model fits a model for the outcome, including this predicted treatment as a covariate. This approach calculates confidence intervals that account for the uncertainty in the predicted treatment that was obtained from the first model.

In the Telecare study, the intention-to-treat effect for differences in average levels of function between the physiotherapy and control groups at 6 months was 4.7 units, with a 95% CI from 1.0 to 8.4 units.[3] Of the 87 participants assigned to the intervention arm, 76 (approximately 87%) had five or more sessions. The complier average causal effect, estimated using a two-stage least squares method, was 5.3 units, with a 95% CI ranging from 2.3 to 8.4 units. This effect is interpreted as the effect of the telephone physiotherapy sessions compared with the existing telephone support service among participants who would undertake at least five sessions of telephone physiotherapy when assigned to do so and only if assigned to do so. The complier average causal effect estimate was approximately equal to the intention-to-treat effect estimate divided by the proportion of compliers in the physiotherapy arm (ie, 4.7 divided by 0.87).

Concluding remarks

Researchers conducting studies where participants depart from their assigned treatment may be tempted to estimate effects such as the per-protocol effect. However, this effect does not have a valid causal interpretation and is not recommended. The complier average causal effect is a preferable measure, estimating the effect of the intervention among those participants who would always receive their assigned treatment. While this group of participants cannot be identified, provided the required assumptions are valid, techniques are available to estimate this effect. This Research Note has considered the simple situation of all-or-nothing compliance: methods for partial compliance are also available.[15] Researchers conducting trials should collect data on the actual treatment that participants receive, any reasons for departure from the allocated treatment, and carefully consider the validity of assumptions when estimating the complier average causal effect.

Footnotes: [a]These assumptions are: no unmeasured confounding of the relationship between the exposure and the outcome; all participants have a non-zero probability of receiving all levels of the exposure; the outcome that was observed for a participant is the same as the potential outcome given their actual level of exposure.

Ethics approval: Nil.
Competing interests: Nil.
Source(s) of support: Nil.
Acknowledgements: Nil.
Provenance: Invited. Peer reviewed.

Correspondence: Jessica Kasza, School of Public Health and Preventive Medicine, Monash University, Australia. Email: jessica.kasza@monash.edu

Jessica Kasza
School of Public Health and Preventive Medicine, Monash University, Australia

References

1. Bennell KL, et al. *BMJ Open*. 2017;7:e014658.
2. Hinman RS, et al. *Ann Intern Med*. 2016;165:381–389.
3. Hinman RS, et al. *Br J Sports Med*. 2020;54:790–797.
4. Herbert R. *J Physiother*. 2020;66:273–277.
5. Williamson EJ, et al. *Respirology*. 2014;19:303–311.
6. VanderWeele T. *Ann Stat*. 2013;41:196–220.
7. Hernán MA, Hernández-Díaz S. *Clin Trials*. 2012;9:48–55.
8. Shrier I, et al. *Clin Trials*. 2014;11:28–37.
9. Hernán MA, Robins JM. *Causal inference: what if*. Boca Raton: Chapman & Hall/CRC; 2020.
10. Frangakis CE, Rubin DB. *Biometrics*. 2002;58:21–29.
11. Herbert RD, et al. *BMC Med Res Methodol*. 2018;18:48.
12. White IR, et al. *Stat Med*. 1999;18:2617–2634.
13. Angrist JD, et al. *JASA*. 1996;91:444–455.
14. Stuart EA, et al. *Prev Sci*. 2008;9:288–298.
15. Jin H, Rubin DB. *J Am Stat Assoc*. 2008;103:101–111.