

A simple principal stratum estimator for failure to initiate treatment

Brennan C Kahan¹, Ian R White¹, Mark Edwards², Michael O Harhay^{3,4}

1 MRC Clinical Trials Unit at UCL, London, UK

2 Department of Anaesthesia, University Hospital Southampton NHS Foundation Trust, Southampton, UK

3 Clinical Trials Methods and Outcomes Lab, PAIR (Palliative and Advanced Illness Research) Center, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA

4 Department of Biostatistics, Epidemiology, and Informatics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA

Correspondence to: Brennan Kahan, b.kahan@ucl.ac.uk

Funding

BCK and IRW are funded by the UK MRC, grants MC_UU_00004/07 and MC_UU_00004/09.

Abstract

A common intercurrent event affecting many trials is when some participants do not begin their assigned treatment. For example, in a trial comparing two different methods for fluid delivery during surgery, some participants may have their surgery cancelled. Similarly, in a double-blind drug trial, some participants may not receive any dose of study medication. The commonly used intention-to-treat analysis preserves the randomisation structure, thus protecting against biases from post-randomisation exclusions. However, it estimates a treatment policy effect (i.e. addresses the question “what is the effect of the intervention, regardless of whether the participant actually begins treatment?”), which may not be the most clinically relevant estimand. A principal stratum approach, estimating the treatment effect in the subpopulation of participants who would initiate treatment (regardless of treatment arm), may be a more clinically relevant estimand for many trials. We show that a simple principal stratum estimator based on a “modified intention-to-treat” population, where participants who experience the intercurrent event are excluded, is unbiased for the principal stratum estimand under certain assumptions that are likely to be plausible in many trials, namely that participants who initiate the intervention under one treatment condition would also do so under the other treatment condition. We provide several examples of trials where this assumption is plausible, and several instances where it is not. We conclude that this simple principal stratum estimator can be a useful strategy for handling failure to initiate treatment.

Key words: estimand, estimator, intercurrent event, treatment initiation, randomised controlled trial, principal stratum, modified intention-to-treat

Background

A common intercurrent event affecting many randomised controlled trials (RCTs) is when some participants do not begin the treatment they were randomised to [1]. For example, the FLO-ELA trial (ISRCTN 14729158) compared the use of a cardiac output monitor against usual care in determining when and how much fluid to give to participants undergoing emergency laparotomy. Due to practical considerations, there is usually a small delay between randomisation and the start of surgery, and a small number of participants had their surgery cancelled after randomisation, either because they deteriorated and were too unwell to undergo surgery, or because the underlying issue resolved itself. Here, a cancelled surgery precludes the participant from initiating treatment, as fluids are no longer relevant for participants who do not undergo their planned surgery. Similar issues may arise in other trials, where participants do not begin their study treatment, or do not receive any dose of study drug.

The intention-to-treat (ITT) strategy is often the recommended analysis approach, regardless of the type of intercurrent events affecting the trial [1-3]. Under ITT, all randomised participants are included in the analysis, and analysed according to their randomised group. A key benefit of ITT is that it preserves the randomisation structure, thus protecting against biases that post-randomisation exclusions can introduce [1, 3-6].

However, with the recent publication of the ICH-E9 addendum on estimands [7], there is growing recognition of the value of first defining the treatment effect we wish to estimate (the estimand), and then choosing an analytical strategy that corresponds to that estimand (the estimator) [8-14]. Hence, an ITT analysis should only be used if it is estimating the desired treatment effect. In the FLO-ELA trial, an ITT analysis would include participants who did not undergo surgery. This corresponds to a *treatment policy* strategy [7], which considers the intercurrent event irrelevant (i.e. the intercurrent event is considered part of the treatment strategy). Thus, using ITT compares two different strategies of fluid delivery during surgery, *regardless of whether participants actually undergo surgery*. This estimand is clearly not of direct clinical interest, so although an ITT analysis preserves the randomisation structure, it does not provide a meaningful treatment effect in this setting.

A more meaningful estimand may be one that uses a *principal stratum* strategy [7, 15, 16], where interest lies in the effect of treatment in the subpopulation of participants who actually undergo surgery and thus are able to initiate treatment. Notably, the principal stratum population is defined as participants who would undergo surgery regardless of their assigned treatment group. This is a more clinically meaningful effect, as it addresses the question “What is the effect of treatment in participants who undergo an emergency laparotomy (and thus can actually receive treatment)?” instead of “What is the effect of treatment regardless of whether participants undergo an emergency laparotomy?”

These types of intercurrent events are sometimes handled in the analysis using a “modified ITT” (mITT) analysis, where participants experiencing the intercurrent event are excluded from the analysis population [1-3, 17]. However, this analysis approach has been criticised because it does not preserve the randomisation structure, and can thus induce bias [1]. In this paper, we show that a simple estimator based on a mITT population can, in fact, be unbiased for the principal stratum estimand under certain assumptions which are plausible in many trials. We provide several trials as examples of where the estimator described above is appropriate, and several trials where it would not be appropriate as the required assumptions are unlikely to be fulfilled.

Principal stratum estimand

We begin by defining the principal stratum estimand in terms of the intercurrent event of interest, which we refer to as “failure to initiate treatment”. Depending on the trial, this could be defined to include any cause of initiation failure (e.g. in a placebo-controlled drug trial where participants do not begin their course of study treatment, regardless of reason), or be defined based on a single cause of initiation failure, with other causes handled using alternate strategies (e.g. in FLO-ELA, where surgery cancellation precludes treatment initiation, but participants may also not initiate for other reasons).

The estimand can be written as:

$$E(Y^{(Z=1)} - Y^{(Z=0)} | I^{(Z=1)} = I^{(Z=0)} = 1) \quad (1)$$

where Z denotes treatment (0=control, 1=intervention), $Y^{(Z=1)}$ and $Y^{(Z=0)}$ denote the participant’s potential outcome under intervention and control respectively, and $I^{(Z=1)}$ and $I^{(Z=0)}$ are indicator variables denoting whether the participant would initiate treatment under intervention and control respectively (where 0 denotes they would not initiate treatment, and 1 denotes they would). This estimand implies the treatment effect in the subpopulation of participants for whom $I^{(Z=1)} = I^{(Z=0)} = 1$ (i.e. who would initiate treatment under either treatment arm) is of interest.

We note that the definition above primarily focusses on how the intercurrent event “failure to initiate treatment” is handled; the other aspects which make up the estimand (population, treatment conditions, endpoint, summary measure, and handling of other intercurrent events) would also need to be specified in order for the estimand to be complete [7].

Principal stratum estimator

Our proposed principal stratum estimator in which participants who experience the intercurrent event are excluded (i.e. using a mITT population) can be written as:

$$E(Y|Z = 1, I_{Z=1} = 1) - E(Y|Z = 0, I_{Z=0} = 1) \quad (2)$$

where Y is the participant’s observed outcome and $I_{Z=1}$ is an indicator denoting whether the participant initiated treatment under $Z = 1$ and similarly for $I_{Z=0}$. Hence, this estimator compares the set of participants in the intervention group initiated the intervention treatment against the set of participants in the control group who initiated the control treatment.

To evaluate the properties of this estimator we begin by outlining the four principal strata defined by the intercurrent event “failure to initiate treatment” (Table 1) [16].

Table 1: The four principal strata defined by the intercurrent event “failure to initiate treatment”.

The estimator excluding participants who do not initiate compares the shaded cells in the intervention column against the shaded cells in the control column.

	Initiate treatment?	
	Control ($Z = 0$)	Intervention ($Z = 1$)
Always initiators	Yes	Yes
Intervention initiators	No	Yes
Control initiators	Yes	No
Never initiators	No	No

The “Always initiators” are participants who would initiate treatment in either treatment group (i.e. for whom $I^{(Z=1)} = I^{(Z=0)} = 1$), and the “never initiators” are participants who would not initiate treatment regardless of treatment group (i.e. $I^{(Z=1)} = I^{(Z=0)} = 0$). The “intervention initiators” are participants who would only initiate treatment in the intervention arm (i.e. $I^{(Z=1)} = 1$ and $I^{(Z=0)} = 0$), and the “control initiators” are participants who only initiate treatment if they are in the control group (i.e. $I^{(Z=1)} = 0$ and $I^{(Z=0)} = 1$).

In a RCT, randomisation can be seen to balance the treatment groups within each stratum in expectation. Hence, if we knew the principal stratum membership of each participant, we could perform the analysis in the stratum of interest and obtain an unbiased estimator based on a randomised comparison. However, in practice we do not know the principal stratum membership, as this requires knowledge of participants’ initiation status under both treatment conditions, but we can only observe a participant’s initiation status under the treatment arm to which they were randomised. For instance, if a participant allocated to the intervention initiated treatment, we have no way of telling whether they are an “Always initiator” or an “Intervention initiator”.

We can see from Table 1 that the estimator described in (2) is not based on a randomised comparison, as it compares different principal strata between the treatment arms (i.e. the analysis population differs between treatment groups) [6]. The intervention arm includes the “always initiators” and the “intervention initiators”, while the control group consists of the “always initiators” and the “control initiators”. Hence, if these two strata (the “intervention initiators” and “control initiators”) differ in terms of their potential outcomes or their size as percentages of the overall population, this estimator will be biased.

Assumptions for unbiasedness

Ensuring estimator (2) is unbiased for the principal stratum estimand in (1) requires a way of identifying the principal stratum membership of participants. This is possible under the assumption that there are no “intervention initiators” and no “control initiators” in the trial, i.e. that *if someone does not initiate treatment in one arm, they would also not initiate treatment in the other arm (and, conversely, if someone does initiate treatment in one arm, they would also initiate treatment in the other arm)*.

This assumption implies there are only two principal strata in the trial, “never initiators” and “always initiators”; hence, if we observe a participant’s initiation status in one treatment arm, we can infer their initiation status had they been allocated to the alternative treatment, and identify the principal stratum to which they belong (i.e. if we observe $I_{Z=1} = 1$ then we know that $I_{Z=0} = 1$ as well, and hence the participant belongs to the “always initiators” group, and vice versa).

Then, estimator (2), which only includes participants who initiated treatment in their assigned arm, will in fact be a comparison within the “always initiators” stratum, which, as discussed above, is based on a randomised comparison (Table 2). As such, this estimator will be unbiased for the principal stratum estimand in (1).

This assumption can be partly assessed, on the basis that if it is true, then the proportion of non-initiators should, on average, be the same across randomised groups. Hence, if there are large discrepancies in the proportion of non-initiators between groups, this may provide evidence that this assumption has been violated. However, even if this proportion is the same across groups, this is

not a guarantee the assumption is true, hence this assumption cannot be fully tested, and therefore must rely primarily on contextual information around its underlying plausibility.

Table 2: The principal stratum estimator under the assumption of no “intervention initiators” and no “control initiators”. The estimator compares the shaded cells in the intervention column against the shaded cells in the control column. Grey indicates that these strata do not exist.

	Initiate treatment?	
	Control (Z = 0)	Intervention (Z = 1)
Always initiators	Yes	Yes
Intervention initiators	No	Yes
Control initiators	Yes	No
Never initiators	No	No

We demonstrate this as follows. First, under consistency assumption [18] (which states that if $Z = z$, then $Y = Y^{(Z=z)}$), we can write $Y|Z = 1$ as $Y^{(Z=1)}$ and $Y|Z = 0$ as $Y^{(Z=0)}$, as well as $I_{Z=1}$ as $I^{(Z=1)}$ and similarly for $I_{Z=0}$. Then, using randomisation, the principal stratum estimator in (2):

$$E(Y|Z = 1, I_{Z=1} = 1) - E(Y|Z = 0, I_{Z=0} = 1)$$

Can be written as:

$$= E(Y^{(Z=1)}|I^{(Z=1)} = 1) - E(Y^{(Z=0)}|I^{(Z=0)} = 1)$$

Then, under the assumption that there are no “intervention initiators” and no “control initiators”, $I^{(Z=1)} = 1$ implies that $I^{(Z=0)} = 1$, and vice versa for $I^{(Z=0)}$, implying that $I^{(Z=1)} = I^{(Z=0)} = 1$.

Then, the estimator becomes:

$$\begin{aligned} & E(Y^{(Z=1)}|I^{(Z=1)} = 1) - E(Y^{(Z=0)}|I^{(Z=0)} = 1) \\ &= E(Y^{(Z=1)}|I^{(Z=1)} = I^{(Z=0)} = 1) - E(Y^{(Z=0)}|I^{(Z=1)} = I^{(Z=0)} = 1) \\ &= E(Y^{(Z=1)} - Y^{(Z=0)}|I^{(Z=1)} = I^{(Z=0)} = 1) \end{aligned}$$

which is the principal stratum estimand in (1).

We note that this estimator is unbiased regardless of what happens to “never initiators” (e.g. regardless of whether they receive no treatment at all, or receive a non-trial treatment instead).

Determining whether the principal stratum estimator is appropriate

Two key factors can be used to determine whether the “Always initiators” principal stratum can be identified, thus ensuring the principal stratum estimator in (2) is appropriate. First, the intercurrent event of interest must be identifiable in each treatment arm (i.e. we must be able to measure which participants in each treatment group experience the intercurrent event).

Secondly, we must be able to assume that the occurrence of the intercurrent event is not affected by treatment allocation (i.e. that if the intercurrent event occurs for a participant under one treatment condition, it would also occur under the other treatment condition). Together, these two factors justify the assumption of no “intervention initiators” and no “control initiators”.

Below we give several examples of trials both where the principal stratum estimator is and is not appropriate (Table 3).

FLO-ELA

As described earlier, FLO-ELA is an open-label trial comparing two methods of fluid delivery (cardiac output monitor vs. clinician judgement) in participants undergoing emergency laparotomy. The primary outcome measure is the number of days that participants are alive and out of hospital within 90 days of randomisation. The intercurrent event of interest is the cancellation of surgery, which precludes treatment from being initiated (as fluid delivery is only relevant once participants begin surgery). This is an example of a “cause-specific” initiation failure, as some participants who do undergo surgery may also not initiate treatment for other reasons. We note that ideally randomisation would be done at the point surgery was due to begin, to minimise the number of cancellations, however in practice this is not always feasible due to the time required to prepare the intervention combined with the emergency setting.

In this trial, the principal stratum estimator is appropriate. First, the intercurrent event can be easily identified and measured in both treatment groups, and secondly, the decision on whether to cancel surgery or not will almost certainly not be affected by the allocated treatment arm. This is because in most cases the relevant decision makers (surgeons) will be unaware of trial group allocation until surgery starts (i.e. at the point the decision is made). Further, the decision not to proceed with surgery has large health implications for the patient, and is only undertaken in response to a major change in the patient’s clinical condition since surgery was initially planned, and it is implausible that such a fundamental change in patient care would be undertaken on the basis of the planned method of fluid delivery.

MIST2

MIST2 was a 2x2 factorial trial evaluating the use of tPA and DNase in participants with pleural infection [19]. We focus on the tPA comparison here. tPA was compared in a double-blind fashion against a matching placebo, so that neither participants nor clinicians were aware of whether participants were receiving tPA or placebo. Treatment was to be given for three days. The intercurrent event of interest is failure to begin treatment (i.e. if participants did not receive any dose of study drug), regardless of the reason.

In this trial, the principal stratum estimator is appropriate. First, the intercurrent event (“failure to begin study treatment”) can be identified and measured in each treatment arm. Secondly, treatment allocation should not influence the occurrence of the intercurrent event, as the trial is double-blinded, so those deciding on whether to initiate treatment will not be aware of whether it is tPA or placebo being initiated.

COPERS

The COPERS trial evaluated the use of a group pain management support intervention for participants with chronic musculoskeletal pain [20]. The intervention consisted of a group intervention delivered over three days with a follow-up session at two weeks. The control consisted of usual care. The intercurrent event of interest is failure to begin the group intervention session (i.e. failure to attend any sessions).

Here, the principal stratum estimator is *not* appropriate, as there is no way of identifying which participants in the control arm would experience the intercurrent event, and so the principal stratum of interest cannot be identified.

SWAP

The SWAP trial evaluated the use of a multi-modal group intervention to reduce weight in obese adults [21]. The intervention consisted of eight group sessions delivered by trained advisors, while the control consisted of four standard weight management sessions delivered by a practice nurse. Due to the nature of the interventions, the trial was open-label, with participants and healthcare professionals aware of the allocated group. The intercurrent event of interest is failure to attend any treatment sessions.

In this trial, the principal stratum estimator is *not* appropriate. Although the intercurrent event can be identified in each treatment arm, the occurrence of the intercurrent event may be affected by the allocated arm. For instance, due to participants’ perceptions about the two treatment arms, they may be more likely to attend the group intervention compared to the standard nurse-led weight management sessions, or vice versa. Hence, the assumption of no “intervention initiators” and no “control initiators” is likely to be false, and the estimator will be based on a non-random comparison, and hence subject to bias.

Table 3 – Examples of trials where the principal stratum estimator both is and is not appropriate

Trial	Description	Intercurrent event	Principal stratum estimator appropriate?	Rationale
FLO-ELA	An open-label trial comparing use of a cardiac output monitor vs. clinical judgement for fluid delivery in participants undergoing emergency laparotomy	Cancellation of surgery	Yes	Intercurrent event (“surgery cancellation”) can be identified in both treatment arms, and the decision to cancel surgery is unlikely to be affected by the arm that participants are allocated to
MIST2	A blinded trial comparing tPA vs. placebo in participants with pleural infection	Failure to receive any dose of study treatment	Yes	Intercurrent event (“failure to initiate treatment”) can be identified in both arms, and will not be affected by the arm that participants are allocated to due to blinding
COPERS	An open-label trial comparing a group pain management support intervention vs. usual care in participants with	Failure to attend any sessions of the group intervention	No	There is no way of identifying which participants in the usual care arm would attend the group intervention,

	chronic musculoskeletal pain			and hence no way of identifying the principal stratum of interest
SWAP	An open-label trial comparing a group intervention vs. a standard weight management session delivered by practice nurses in obese adults	Failure to attend any sessions of allocated treatment	No	Attendance may be affected by the allocated treatment arm (e.g. participants may be more likely to attend the group intervention compared to the standard weight management sessions, or vice versa), and thus the estimator will be based on a non-randomised comparison

Reporting of this estimator

Given that the principal stratum estimator in (2) relies on certain assumptions for unbiasedness, its use must be transparently reported so that readers may evaluate whether the assumptions underpinning the study results are reasonable.

To provide full clarity, we suggest the methods be reported according to Box 1. This provides the target of estimation (the estimand), which alerts readers a principal stratum effect is of interest; explicitly states that participants who experience the intercurrent event are excluded from the analysis population, along with the reason for doing so; provides the assumptions required for this estimator to be unbiased; and finally, justifies those assumptions for the specific trial. Together, this provides readers with an understanding of *what* is being estimated, how the estimation procedure works, the assumptions required, as well as the rationale underpinning those assumptions.

Box 1 – Reporting recommendations for using the principal stratum estimator. Recommendations are based on the principal stratum estimator defined in formula (2) for the intercurrent event “failure to initiate treatment”.

- In the estimand definition, state that a principal stratum approach is being used for the intercurrent event
- In the description of the estimator, state that participants who experience the intercurrent event are excluded from the analysis population as a way to estimate the principal stratum effect
- State the assumptions required for this estimator to be unbiased (i.e. no “intervention initiators” and no “control initiators”)
- Provide an explanation for why the assumptions above are justified in the trial

Discussion

Failure to initiate treatment is a common intercurrent event in randomised trials. In this paper, we argue that a simple estimator which excludes participants who experience the event is unbiased for

a principal stratum estimand under the assumption that if a participant experiences the intercurrent event under one treatment condition, they would also experience it under the other conditions.

This assumption is likely to be fulfilled in many settings, including most double-blind trials, where the decision to initiate treatment will not be affected by the treatment arm. However, this assumption may also be plausible in many open-label trials as well; for instance, in FLO-ELA, there is strong justification for expecting cancellation of surgery not to be affected by treatment group.

In many settings a principal stratum effect may be more clinically relevant and generalisable than the treatment policy approach for these types of intercurrent events. For example, in the FLO-ELA trial, cancellation of surgery renders the participant no longer part of the population of interest (which is participants undergoing emergency laparotomy); hence, a principal stratum approach may be more reflective of the treatment effect in routine clinical practice compared to a treatment policy approach.

Modified ITT analyses have been criticised on the basis of ambiguity, i.e. the term has been used by different researchers to mean different things [2, 3, 5, 17]. If using this term to describe the estimator, we recommend it be described as in Box 1 (i.e. that it is explicitly stated which intercurrent events are excluded), to avoid potential misunderstandings. It is also not always clear which estimand such mITT analyses have targeted in the past. The results in this paper can serve to clarify the target of estimation for these analyses, as well as the assumptions required to justify their validity.

Although this paper has focussed on the intercurrent event “failure to initiate treatment”, the results could also apply to other intercurrent events where the assumptions discussed here are likely to be fulfilled. For instance, in trials where mortality is an intercurrent event but the treatments are known to not affect mortality (for instance, in certain palliative care trials), then (if clinically appropriate) the principal stratum estimator described in (2) could also be used.

Conclusion

A simple principal stratum estimator which excludes participants who experience the intercurrent event can be a useful way to estimate a principal stratum effect, provided the assumptions are justified.

Funding

BCK, and IRW are funded by the UK MRC, grants MC_UU_00004/07 and MC_UU_00004/09.

Contributions

BCK wrote the first draft of the manuscript. IRW, ME, and MH revised the manuscript. All authors read and approved the final manuscript.

Declaration of Conflicting Interests

The Authors declare that there is no conflict of interest.

Data availability

This article contains no data.

References

1. Moher D, Hopewell S, Schulz KF, Montori V, Gotzsche PC, Devereaux PJ, et al. CONSORT 2010 explanation and elaboration: updated guidelines for reporting parallel group randomised trials. *BMJ*. 2010;340:c869.
2. Dodd S, White IR, Williamson P. Nonadherence to treatment protocol in published randomised controlled trials: a review. *Trials*. 2012;13:84.
3. Joseph R, Sim J, Ogollah R, Lewis M. A systematic review finds variable use of the intention-to-treat principle in musculoskeletal randomized controlled trials with missing data. *J Clin Epidemiol*. 2015;68(1):15-24.
4. Detry MA, Lewis RJ. The intention-to-treat principle: how to assess the true effect of choosing a medical treatment. *JAMA*. 2014;312(1):85-6.
5. Hollis S, Campbell F. What is meant by intention to treat analysis? Survey of published randomised controlled trials. *BMJ*. 1999;319(7211):670-4.
6. Abraha I, Montedori A. Modified intention to treat reporting in randomised controlled trials: systematic review. *Bmj*. 2010;340:c2697.
7. ICH E9 (R1) addendum on estimands and sensitivity analysis in clinical trials to the guideline on statistical principles for clinical trials [Available from: https://www.ema.europa.eu/en/documents/scientific-guideline/ich-e9-r1-addendum-estimands-sensitivity-analysis-clinical-trials-guideline-statistical-principles_en.pdf].
8. Clark TP, Kahan BC, Phillips A, White I, Carpenter JR. Estimands: bringing clarity and focus to research questions in clinical trials. *BMJ Open*. 2022;12(1):e052953.
9. Degtyarev E, Rufibach K, Shentu Y, Yung G, Casey M, Englert S, et al. Assessing the Impact of COVID-19 on the Clinical Trial Objective and Analysis of Oncology Clinical Trials-Application of the Estimand Framework. *Stat Biopharm Res*. 2020;12(4):427-37.
10. Fiero MH, Pe M, Weinstock C, King-Kallimanis BL, Komo S, Klepin HD, et al. Demystifying the estimand framework: a case study using patient-reported outcomes in oncology. *Lancet Oncol*. 2020;21(10):e488-e94.
11. Kahan BC, Morris TP, White IR, Carpenter J, Cro S. Estimands in published protocols of randomised trials: urgent improvement needed. *Trials*. 2021;22(1):686.
12. Kahan BC, Morris TP, White IR, Tweed CD, Cro S, Dahly D, et al. Treatment estimands in clinical trials of patients hospitalised for COVID-19: ensuring trials ask the right questions. *BMC Med*. 2020;18(1):286.
13. Lawrance R, Degtyarev E, Griffiths P, Trask P, Lau H, D'Alessio D, et al. What is an estimand & how does it relate to quantifying the effect of treatment on patient-reported quality of life outcomes in clinical trials? *Journal of Patient-Reported Outcomes*. 2020;4(1):68.
14. Sun S, Weber HJ, Butler E, Rufibach K, Roychoudhury S. Estimands in hematologic oncology trials. *Pharm Stat*. 2021;20(4):793-805.
15. Bornkamp B, Rufibach K, Lin J, Liu Y, Mehrotra DV, Roychoudhury S, et al. Principal stratum strategy: Potential role in drug development. *Pharm Stat*. 2021;20(4):737-51.
16. Frangakis CE, Rubin DB. Principal stratification in causal inference. *Biometrics*. 2002;58(1):21-9.
17. Rehal S, Morris TP, Fielding K, Carpenter JR, Phillips PPJ. Non-inferiority trials: are they inferior? A systematic review of reporting in major medical journals. *BMJ Open*. 2016;6(10):e012594.
18. Hernán MA, Robins JM. *Causal Inference: What If*. Boca Raton: Chapman & Hall/CRC; 2020.
19. Rahman NM, Maskell NA, West A, Teoh R, Arnold A, Mackinlay C, et al. Intrapleural use of tissue plasminogen activator and DNase in pleural infection. *N Engl J Med*. 2011;365(6):518-26.
20. Taylor SJ, Carnes D, Homer K, Kahan BC, Hounscome N, Eldridge S, et al. Novel Three-Day, Community-Based, Nonpharmacological Group Intervention for Chronic Musculoskeletal Pain (COPERS): A Randomised Clinical Trial. *PLoS Med*. 2016;13(6):e1002040.

21. McRobbie H, Hajek P, Peerbux S, Kahan BC, Eldridge S, Trepel D, et al. Randomised controlled trial and economic evaluation of a task-based weight management group programme. *BMC Public Health*. 2019;19(1):365.