International Journal of Epidemiology, 2017, 96–102
doi: 10.1093/ije/dyw050
Advance Access Publication Date: 25 April 2016
Original article

Methodology

The cohort multiple randomized controlled trial design: a valid and efficient alternative to pragmatic trials?

Joanne M van der Velden,[1]* Helena M Verkooijen,[1,2] Danny A Young-Afat,[1,2] Johannes PM Burbach,[1] Marco van Vulpen,[1] Clare Relton,[3,4] Carla H van Gils,[2] Anne M May[2] and Rolf HH Groenwold[2]

[1]Department of Radiation Oncology, [2]Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, Utrecht, The Netherlands, [3]School for Health and Related Research, University of Sheffield, Sheffield, UK and [4]School of Healthcare, University of Leeds, Leeds, UK

*Corresponding author. Department of Radiation Oncology, Imaging Division, University Medical Center Utrecht, P.O. Box 85500, 3508 GA Utrecht, The Netherlands. E-mail: j.m.vandervelden@umcutrecht.nl.

Accepted 23 February 2016

Abstract

Randomized controlled trials (RCTs)—the gold standard for evaluating the effects of medical interventions—are notoriously challenging in terms of logistics, planning and costs. The cohort multiple randomized controlled trial approach is designed to facilitate randomized trials for pragmatic evaluation of (new) interventions and is a promising variation from conventional pragmatic RCTs. In this paper, we evaluate methodological challenges of conducting an RCT within a cohort. We argue that equally valid results can be obtained from trials conducted within cohorts as from pragmatic RCTs. However, whether this design is more efficient compared with conducting a pragmatic RCT depends on the amount and nature of non-compliance in the intervention arm.

Key words: Pragmatic randomized controlled trials, cohort multiple randomized controlled trial, validity, efficiency, instrumental variable analysis

Key messages

- The cohort multiple randomized controlled trial design was proposed as a variation of the pragmatic randomized controlled trial (RCT). A major difference between an RCT conducted within a cohort (cmRCT) and a pragmatic RCT is the timing of randomization relative to informed consent: in a cmRCT, only participants allocated to the intervention arm are asked for informed consent, which happens after they have been randomized to the intervention.
- Due to timing of the randomization, the non-compliance in the intervention arm may be higher in a cmRCT compared with an RCT. An increased rate of non-compliance in the intervention arm may be (partially) compensated for by lower (or even absent) non-compliance in the control arm. Non-compliance in the intervention arm can be accounted for by instrumental variable analysis.
- Participants allocated to the control arm of a cmRCT are unaware of being in the control group. This may reduce the risk of cross-over, drop-out and reporting bias. As compared with RCTs, standard of care applied in cmRCTs will better resemble routine care.
- More research is needed regarding generalizability of trial results, misclassification of the actual intervention status and implications of conducting multiple RCTs within a cohort.

Introduction

Randomized controlled trials (RCTs) are considered the gold standard for evaluating the effects of medical interventions. Pragmatic trials investigate the effectiveness of medical interventions or strategies under usual conditions.[1,2] In contrast to explanatory trials, these trials are not placebo-controlled and typically do not blind the participants. Results of pragmatic trials often better reflect the effects to be expected in daily practice.[2] Both explanatory and pragmatic trials are notoriously challenging in terms of logistics, planning and costs. Less than one-third of RCTs achieve their planned recruitment target and follow-up is labour intensive.[3,4] What is more, pragmatic RCTs, where participants are not blinded for the intervention status, may be complicated by response bias (also referred to as disappointment bias), and a considerable risk of non-compliance and cross-over between study arms.[5,6]

To overcome these challenges, the 'cohort multiple randomized controlled trial' design was proposed as a variation of pragmatic RCTs.[7] The basis of this design is a prospective cohort of participants with the condition of interest, receiving care as usual, who give informed consent for cohort participation. In our centre, participants are furthermore asked for informed consent to be randomized in future RCTs conducted within the cohort. Participants are informed that they will be offered the experimental intervention if they are randomly selected. They are also informed that they otherwise might serve as controls without being notified and that their data can be used in a trial context.[8] For each participant in the cohort, clinical and patient-reported outcomes are captured at baseline and at regular intervals during follow-up. Within this cohort, multiple RCTs can be conducted. For this purpose, eligible participants who have provided the consent required for them to participate in an RCT within the cohort are identified. From this subcohort, a random selection of participants will be invited to undergo the experimental intervention. Eligible participants who were not randomly selected receive standard care, are not informed about the experimental intervention and serve as controls. Outcomes in this control group are compared with the outcomes of those who were offered the experimental intervention, in order to estimate the effect of the experimental intervention versus usual care. Within the cohort, the same process can be repeated for trials of other interventions. The design appears especially attractive for clinical research areas where many interventions need evaluation, and for highly desired or expensive interventions.[7]

The cohort multiple randomized controlled trial design is gaining interest in different fields of research.[9–16] Given the novelty of the design, several ethical and methodological aspects need in-depth evaluation. Ethical issues have been described elsewhere.[8] In this article we focus on the methodological issues of conducting one RCT within a cohort (cmRCT). We compare cmRCTs and pragmatic RCTs in terms of validity of the results and discuss approaches for analysis of a cmRCT.

Validity of cmRCT results

To obtain a valid estimate of the intervention effect, the group receiving the experimental intervention and the group receiving the standard intervention need to be comparable at the start of the study, during follow-up and at the end of the study.[17] However, at each of these moments, differences between cmRCTs and RCTs may occur.

At the start of the study: timing of randomization

Comparability of intervention groups at the start of the study is most effectively achieved by randomization. A

major difference between a cmRCT and an RCT is the timing of randomization relative to the informed consent procedure. In an RCT, all participants are randomized after they have been informed about the intervention and after they consented to participate in the trial. In a cmRCT however, only participants allocated to the experimental intervention arm are informed about the intervention, but only after they have been randomized. Consent to participate is only sought from participants who are randomized to the experimental arm. This pre-randomization is different from, for example, the Zelen design. In the Zelen design, participants are randomized before seeking consent,[18] whereas participants in a cmRCT have given informed consent to be randomized, although the intervention is not known yet. Participants who are randomized to the experimental arm, may subsequently decline the experimental intervention (non-compliers). Here, the term non-compliance is used to indicate that participants who are allocated to one intervention arm decline that particular intervention at baseline. As a result, the proportion of non-compliance in the experimental arm is expected to be higher in a cmRCT compared with an RCT. This may particularly affect trials looking at interventions that are unpopular among participants, for example time-consuming or inconvenient interventions. In case of non-compliance, per protocol analysis may result in biased estimates of the intervention effect, if reasons for non-compliance are related to the outcome.[19] In RCTs testing an inconvenient, unpopular intervention, many participants will refuse to participate. Only a small subset of eligible participants may be willing to participate, possibly impairing generalizability. In a cmRCT, all eligible participants will be randomized, but participants allocated to the experimental intervention arm will be more likely to refuse the intervention. Hence, trial results will be generalizable to a broader population, but compared with an RCT, intention-to-treat (ITT) analysis will provide a more diluted estimate of the true intervention effect.

Challenges during follow-up

In a randomized double-blind placebo controlled-trial, blinding is relatively straightforward by using a placebo intervention for the control group. Because participants and their physicians are blinded, comparability of the intervention groups during follow-up is likely to be maintained. In contrast, pragmatic RCTs compare interventions under usual conditions, thus participants are not blinded and changes in, for example, health-related behaviour may differ between study arms.[20] Furthermore, participants may drop out if they are not allocated to the intervention they had hoped for[21] or cross over to the preferred intervention arm, especially if the intervention is widely accessible to participants, such as exercise programmes. Participants in cmRCTs are not blinded, but participants in the control group are unaware of being in the control group of a specific trial. As a result, standard of care will not be affected by intervention allocation and will better resemble routine standard of care. Furthermore, drop-out rates may be lower in cmRCTs, since participants in the control group are not likely to withdraw from standard care. Moreover, information on baseline characteristics and outcome measurements (i.e. the regular cohort measures) of drop-outs are still recorded, which is essential in the data analysis.

Measurement of endpoints

Ideally, the assessor of the outcome in trials is blinded for intervention status in order to prevent observer effects.[17] In pragmatic RCTs as well as in cmRCTs, participants are not blinded, which may lead to differential reporting of outcomes particularly in case of patient-reported outcomes. Some participants will consent to participate in a pragmatic RCT because they wish to receive the experimental intervention. They may be disappointed if allocated to the control arm and may subjectively report worse outcomes than were actually experienced, which may bias the observed differences in outcome between interventions.[6,12] This is unlikely to happen among control participants in a cmRCT, since they do not know that they serve as control participants; this leaves potential bias in reported outcomes of patients in the experimental arm (i.e. probably better outcomes than were experienced) only. Therefore in comparison with pragmatic RCTs, the potential for reporting bias may be reduced in cmRCTs.

Analysis of a cmRCT

The primary analysis in an RCT is typically an intention-to-treat (ITT) analysis, which maintains baseline comparability achieved by randomization.[2] Usually, in RCTs with blinded participants, compliance is high. In pragmatic trials, however, it is difficult and often undesirable to blind participants. This increases the risk of non-compliance, leading to underestimation of the true effect (i.e. the effect that would be observed under perfect compliance) in ITT analysis. In pragmatic trials investigating the effects of an intervention under usual conditions, non-compliance can be seen as part of the intervention effect. However, researchers still might be interested in the 'explanatory' or 'real' effect of the intervention under perfect conditions. To control for non-compliance in

RCTs, instrumental variable (IV) analysis (or Complier Average Causal Effect (CACE) analysis) may be used to account for non-compliance.[19,22,23] The IV analysis accounts for non-compliance by inflating the ITT effect to the effect that would be observed in the (possibly hypothetical) situation of perfect compliance. The estimated effect applies to those who comply with the offered intervention (Box 1).

Box 1 IV analysis in an RCT

Consider an RCT with Z as an indicator of (random) assignment of the intervention (e.g. experimental intervention = 1, usual care = 0), X as actual intervention received (e.g. experimental intervention = 1, usual care = 0) and Y as outcome.

$$Z \rightarrow X \rightarrow Y$$

The variable Z is referred to as the instrument, or the instrumental variable. The ITT effect (i.e. the average causal effect of Z on Y) differs from the average causal effect of X on Y if some participants do not comply with the assigned intervention. The smaller the rate of compliance (i.e. the weaker the relation between Z and X), the more the ITT effect will tend to differ from the average causal effect. To obtain the effect that would be observed under perfect compliance (IV effect), the ITT effect needs to be inflated. In IV analysis the average effect of X on Y is estimated from two effects of Z, namely the average effect of Z on Y and the average effect of Z on X in the following way:

$$\frac{Z \rightarrow Y \ \ (\text{ITT effect})}{Z \rightarrow X \ \ (\text{compliance})}$$

To obtain the average intervention effect, one inflates the ITT effect in the numerator of the estimator by dividing by a factor which is lower as compliance decreases. The weaker the association between Z and X, the more the ITT effect will be inflated because of the shrinking denominator. If compliance is perfect (i.e. Z equals X and $Z \rightarrow X = 1$), the ITT effect equals the IV effect. Compliance (i.e. $Z \rightarrow X$) can be estimated as the difference in the observed probabilities of receiving the experimental intervention between the two allocation groups. IV analysis estimates an effect that applies to those who comply with the allocated intervention.[22–24]

ITT versus IV analysis

In a cmRCT, the compliance in the control group (i.e. usual care) will approximate to 100%, since participation in the cohort is conditional on receiving the standard of care. Since control participants are not informed about the experimental intervention, cross-over to the experimental intervention arm is unlikely. Compliance in the experimental intervention arm, however, may be substantially lower than 100%, since participants are free to accept or to decline the experimental intervention. To illustrate the impact of compliance, we compared ITT and IV analysis in both cmRCTs and RCTs. We considered a hypothetical randomized trial with two intervention arms, which is designed to detect a 10% difference in the risk of the outcome. The risk of the outcome is 10% in the experimental intervention arm, and 20% in the control arm. We simulated four approaches: (i) ITT analysis of an RCT; (ii) ITT analysis of a cmRCT; (iii) IV analysis of a RCT; and (iv) IV analysis of a cmRCT (Figure 1). In all four approaches, the true intervention effect is observed when there is perfect compliance. In both RCTs and cmRCTs, the observed risk difference (obtained with ITT analysis) obviously depends on the proportion of non-compliance: the ITT estimate becomes more diluted as non-compliance increases. Since non-compliance occurs in the experimental intervention group only, the dilution is less in the cmRCT scenario, yielding a less biased estimate of the true treatment effect. Due to the timing of randomization in a cmRCT—consent to participate is sought after randomization—higher non-compliance is to be expected in the experimental arm of the cmRCT compared with the experimental arm of a pragmatic RCT. However, because of the high compliance in the control arm, a cmRCT has room for more non-compliance in the experimental arm in comparison with an RCT in which non-compliance is expected in both intervention arms (Figure 1a, b). Possibly there will be situations in which the amount of non-compliance in the experimental arm is too large to be compensated for by the low amount of non-compliance in the control arm.

The intervention effect among compliers is estimated using the IV analysis, at the expense of precision. This imprecision increases more slowly in the cmRCT scenario since the probability of receiving the experimental intervention in the control group is very low and non-compliance is in the experimental intervention arm only. Note that at the same total amount of non-compliance—thus a double amount of non-compliance in the experimental intervention and zero non-compliance in the control arm—the results from a cmRCT are comparable to the results from an RCT (Figure 1c, d). Still, the main advantage of a cmRCT is the containment and control of the non-

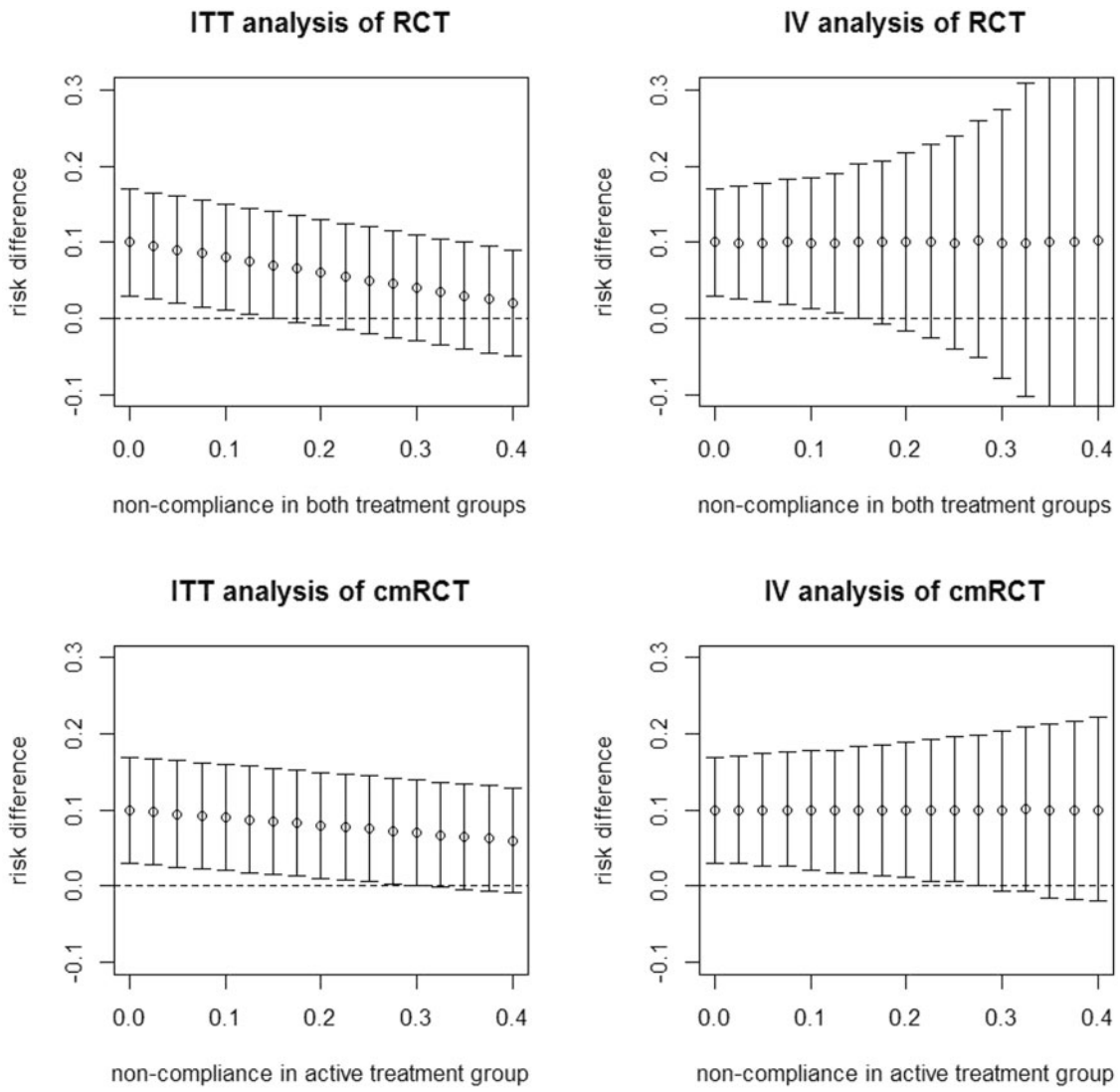

Figure 1. Numerical example of ITT analysis and IV analysis in both RCTs and cmRCT studies. Note that non-compliance in RCTs is possible in both intervention arms whereas, in cmRCT studies, non-compliance will be in the experimental treatment arm only. In a RCT at 10% non-compliance, there is 10% non-compliance in both arms, resulting in a 10% total non-compliance. This is in contrast to 10% non-compliance in a cmRCT, since non-compliance occurs in the experimental intervention arm only. This results in a 5% total non-compliance. Higher non-compliance is to be expected in the experimental intervention arm of the cmRCT. However, a cmRCT has room for more non-compliance in the experimental intervention group in comparison with an RCT in which non-compliance is expected in both intervention arms.

compliance, since all non-compliance accumulates in the experimental arm. Conducting a cmRCT, one should consider the amount of non-compliance to be expected in the experimental arm and the accessibility to this experimental intervention.

Assumptions of IV analysis

In RCTs, the indicator of (random) assignment of the intervention can be considered as an IV and thus used in IV analysis in order to estimate the average causal effect of the intervention. An IV should satisfy three key assumptions: the IV is predictive of actual intervention status, does not share common causes with the outcome and affects the outcome only through the intervention.[22–24] The first and second assumptions easily hold in RCTs and cmRCTs, since allocation of the intervention (i.e. the IV) likely will be associated with actual intervention status, and by randomization assumption two is met as well. However, assumption three may be violated in pragmatic RCTs where participants are deliberately not blinded. In these cases, participants may change their behaviour when (not being) offered the intervention. Therefore, random allocation may affect the outcome via, for example, lifestyle changes as well as via the intervention. In a cmRCT however, participants allocated to the control group do not know that they serve as controls. Therefore, assumption three might be less violated in cmRCTs than in pragmatic RCTs.

Discussion

In this paper, we compared cmRCTs with pragmatic RCTs and explored approaches to analyse cmRCT results. Participants in a cmRCT are recruited from an underlying cohort and outcomes measured in this cohort are relevant for the RCTs conducted within that cohort. Therefore, this design would be mostly applicable in cohorts specifically designed as cmRCT cohorts. Once such a cohort has been established, setting up trials will likely be less expensive and will require less effort compared with RCTs because a research infrastructure is already in place. Moreover, the cohort allows for unequal randomization by making use of the (large) control group of the cohort. This may be especially attractive in the case of expensive experimental treatments, to reduce the costs of a trial. Another advantage of the cmRCT is that participants in the control group are unaware that they are participating as controls in a randomized trial. This will reduce not only the potential of reporting bias, but also cross-over of participants from control arm to experimental treatment arm.

Because of its design, cmRCTs are most suitable to evaluate experimental interventions that are not easily accessible for participants. If the intervention under study is in fact accessible to those in the control group (i.e. usual care), compliance in that group may be less than 100% since participants may undergo the experimental treatment on their own initiative. For example, a cmRCT studying the relative effectiveness of two pharmacological drugs that are already on the market will face this challenge; yet this seems unlikely when comparing drugs in a pre-licensing stage. However, to emphasize again, one of the advantages of cmRCTs is that by not informing the control group, contamination may be limited.

So far, we (implicitly) discussed cmRCTs in the context of studies assessing superiority of one intervention over another. Alternatively, the aim of a trial might be to show non-inferiority or equivalence of two interventions. In non-inferiority and equivalence trials, an ITT analysis is anti-conservative,[25] particularly when non-compliance rates are high. IV analysis, as applied in the analysis of a cmRCT, may partly overcome this problem. Note that precision of IV estimates will be smaller (i.e. wider confidence intervals) than the precision of estimates from an ITT analysis conducted in a study with full compliance.

Very few trials are purely explanatory or pragmatic; there is a continuum rather than dichotomy.[2] Different choices in design result in a more pragmatic or more explanatory trial, for example design choices described in the PRECIS-2 tool such as eligibility criteria, setting of a trial and follow-up.[2] In an RCT, randomization is the essential feature; all other design features are optional. The choices made regarding these other design options will make a trial more pragmatic or more explanatory. By design, the comparator in a cmRCT will always be care as usual, making use of existing staff and resources. This is extremely pragmatic in nature. Participants are recruited from an underlying cohort in which all participants with the condition of interest and receiving usual care are enrolled. It is considered very pragmatic to recruit patients in usual care without overt recruitment effort.[2] Moreover, the cmRCT participants recruited from a cohort may also better resemble the population of (future) users of the intervention under study, which again can be considered pragmatic.[2] However, more explanatory choices could be made as well. Very tight selection criteria could still be applied, resulting in a more explanatory cmRCT. The cohort provides regular outcome measurements but presumably more than are done in usual practice. Various adjustments to the intensity of these measurements will move the trial toward the explanatory end of the continuum. Specific directions for administering the experimental intervention by practitioners deemed to have sufficient experience will also result in a more explanatory trial. Just like an RCT, a cmRCT will not automatically answer a purely pragmatic research question since several explanatory features may be included in the design. However, since the comparator will always be care as usual and participants are recruited from an underlying cohort, all cmRCTs are likely to be located at the pragmatic end of the pragmatic–explanatory continuum.

Conclusion

A major difference between an RCT and an RCT within a cohort (cmRCT) is the timing of randomization. Participants in an RCT are randomized to intervention arms after they consent to participation. This is in contrast to a cmRCT in which only participants allocated to the experimental arm are asked for consent to receive the intervention, and only after they have been randomized. Therefore, non-compliance in the experimental arm may be higher in a cmRCT compared with an RCT. On the other hand, control participants in a cmRCT do not know they are in the control arm, which will better mimic routine standard of care and lower the risk of loss to follow-up and response bias. Future studies implementing the cohort multiple randomized controlled trial design need to be conducted in order to quantify the magnitude of these phenomena. Based on our evaluation, we conclude that results from single cmRCTs are as valid as those from pragmatic RCTs. Whether the cohort multiple randomized controlled trial design is more efficient compared with pragmatic RCTs depends on the amount and nature of non-compliance.

Funding

This work was supported by The Netherlands Organization for Scientific Research [916.13.028 to R.H.H. and 016.156.050 to A.M.M.].

Conflict of interest: None.

References

1. Schwartz D, Lellouch J. Explanatory and pragmatic attitudes in therapeutical trials. *J Chron Dis* 1967;**20**:637–48.
2. Loudon K, Treweek S, Sullivan F, Donnan P, Thorpe KE, Zwarenstein M. The PRECIS-2 tool: designing trials that are fit for purpose. *BMJ* 2015;**350**:h2147.
3. McDonald AM, Knight RC, Campbell MK *et al*. What influences recruitment to randomized controlled trials? A review of trials funded by two UK funding agencies. *Trials* 2006;**7**:9.
4. Young RC. Cancer clinical trials – a chronic but curable crisis. *N Engl J Med* 2010;**363**:306–09.
5. Homer CS. Using the Zelen design in randomized controlled trials: debates and controversies. *J Adv Nurs* 2002;**38**:200–07.
6. Sedgwick P. Explanatory trials versus pragmatic trials. *BMJ* 2014;**349**:g6694.
7. Relton C, Torgerson D, O'Cathain A, Nicholl J. Rethinking pragmatic randomized controlled trials: introducing the "cohort multiple randomized controlled trial" design. *BMJ* 2010; **340**:c1066.
8. Young-Afat DA, Verkooijen HM, Van Gils CH *et al*. Staged-informed consent in the cohort multiple randomizedcontrolled trial design. Epidemiology 2016, Jan 6. PMID: 26745611. [Epub ahead of print.]
9. Relton C, Bissell P, Smith C *et al*. South Yorkshire Cohort: a 'cohort trials facility' study of health and weight. Protocol for the recruitment phase. *BMC Public Health* 2011;**11**:640.
10. Griffin XL, Achten J, Parsons N, Boardman F, Griffiths F, Costa ML. The Warwick Hip Trauma Evaluation – an abridged protocol for the WHiTE Study: A multiple embedded randomized controlled trial cohort study. *Bone Joint Res* 2012;**1**:310–14.
11. Peckham EJ, Relton C, Raw J, Walters C, Thomas K, Smith C. A protocol for a trial of homeopathic intervention for irritable bowel syndrome. *BMC Complement Altern Med* 2012;**12**:212.
12. Kwakkenbos L, Jewett LR, Baron M *et al*. The Scleroderma Participant-centered Intervention Network (SPIN) Cohort: protocol for a cohort multiple randomized controlled trial (cmRCT) design to support trials of psychosocial and rehabilitation interventions in a rare disease context. *BMJ Open* 2013;**3**:e003563.
13. Uher R, Cumby J, MacKenzie LE *et al*. A familial risk enriched cohort as a platform for testing early interventions to prevent severe mental illness. *BMC Psychiatry* 2014;**14**:344.
14. Viksveen P, Relton C. Depression treated by homeopaths: a study protocol for a pragmatic cohort multiple randomized controlled trial. *Homeopathy* 2014;**103**:147–52.
15. Dascanio V, Birks Y, Clark L, Fairhurst C, MacPherson H, Torgerson DJ. Randomized cohort trial was shown to be feasible for evaluating treatments in low back pain. *J Clin Epidemiol* 2014;**67**:940–46.
16. Burbach JP, Verkooijen HM, Intven M *et al*. Randomized controlled trial for pre-operative dose-escalation BOOST in locally advanced rectal cancer (RECTAL BOOST study): study protocol for a randomized controlled trial. *Trials* 2015;**16**:58.
17. Grobbee DE, Hoes AW. *Clinical Epidemiology: Principles, Methods and Applications for Clinical Research*. 2nd edn. Sudbury, UK: Jones and Bartlett Publishers, 2014.
18. Zelen M. A new design for randomized clinical trials. Beyond the intention-to-treat in comparative effectiveness research. *N Engl J Med* 1979;**300**:1242–45.
19. Sussman JB, Hayward RA. An IV for the RCT: using instrumental variables to adjust for intervention contamination in randomized controlled trials. *BMJ* 2010;**340**:c2073.
20. Travier N, Velthuis MJ, Steins Bisschop CN *et al*. Effects of an 18-week exercise programme started early during breast cancer intervention: a randomized controlled trial. *BMC Med* 2015;**13**:121.
21. Lindström D, Sundberg-Petersson I, Adami J, Tönnesen H. Disappointment and drop-out rate after being allocated to control group in a smoking cessation trial. *Contemp Clin Trials* 2010;**31**:22–26.
22. Hernan MA, Robbins JM. Instruments for causal inference an epidemiologist's dream? *Epidemiology* 2006;**17**:360–72.
23. Martens EP, Pestman WR, de Boer A, Belitser SV, Klungel OH. Instrumental variables: application and limitations. *Epidemiology* 2006;**17**:260–67.
24. Greenland S. An introduction to instrumental variables for epidemiologists. *Int J Epidemiol* 2000;**29**:722–29.
25. Hernán MA, Hernández-Díaz S. Beyond the intention-to-treat in comparative effectiveness research. *Trials* 2012;**9**:48–55.