



HHS Public Access

Author manuscript

J Pain Symptom Manage. Author manuscript; available in PMC 2016 March 01.

Published in final edited form as:

J Pain Symptom Manage. 2015 March ; 49(3): 625–631. doi:10.1016/j.jpainsymman.2014.11.294.

Which Treatment Is Better? Ascertaining Patient Preferences with Crossover Randomized Controlled Trials

David Hui, MD, MSc, Donna S Zhukovsky, MD, and Eduardo Bruera, MD

Department of Palliative Care and Rehabilitation Medicine, University of Texas M.D. Anderson Cancer Center, Houston, Texas, USA

Abstract

Context—The difference in patient-reported outcomes between study arms can often be difficult to ascertain in randomized controlled trials (RCTs) using a parallel design because of wide inter-individual variations in baseline characteristics and how patients interpret the outcome measures. Furthermore, the minimal clinically important difference is often not available for many outcomes, and even when available, not individualized for each patient. Crossover RCTs are designed for intra-individual comparisons, which can address these issues by asking patients 1) to directly compare the interventions in regard to effectiveness, adverse effects and ease of use, and 2) to provide an overall choice.

Objectives—We discuss the key design elements for crossover trials, their advantages and disadvantages relative to parallel designs, and their utility in palliative care research using a number of case examples.

Methods—This is a narrative review.

Results—Crossover studies randomize patients to a sequence of treatments. In addition to facilitating intra-individual comparisons, they often require a smaller sample size for the same statistical power compared to parallel designs, and are thus less costly. However, crossover studies are only feasible when the condition being studied is relatively stable and the intervention has a short term effect. Crossover studies with inadequate washout periods may be difficult to interpret. The risk of attrition also may increase because of prolonged study duration.

Conclusion—By facilitating intra-individual comparisons and eliciting patient preferences, crossover studies can provide unique information on the superior intervention. Crossover designs should be considered for selected palliative care studies.

© 2014 U.S. Cancer Pain Relief Committee. Elsevier Inc. All rights reserved.

Address correspondence to: David Hui, MD, MSc, Department of Palliative Care & Rehabilitation Medicine Unit 1414, University of Texas M. D. Anderson Cancer Center, 1515 Holcombe Boulevard, Houston, TX, USA 77030, dhui@mdanderson.org.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Disclosures: The authors declare no relevant financial conflicts of interest.

Keywords

clinical trials; crossover studies; palliative care; randomized controlled trial; research design; statistical data interpretation

Introduction

In the era of evidence-based medicine, randomized controlled trials (RCTs) are considered the gold standard to inform clinical practice because they help to minimize selection bias and ascertainment bias. A large majority of RCTs in supportive/palliative care are parallel in design, in which patients are randomized to receive one of the study interventions (e.g., active intervention(s) or control) (1, 2). The primary and/or secondary outcome measures of many supportive/palliative care trials involve patient-reported outcomes (PROs), such as pain, fatigue, quality of life, satisfaction, and preferences. In regard to ascertainment of outcomes for RCTs, there are three important questions:

1. Are there any statistically significant benefits and risks associated with the active intervention?
2. If yes, is the magnitude of the benefits and risks clinically meaningful?
3. What is the overall patient preference, taking into account both the risks and benefits? This information is particularly useful if the choice is made in a blinded fashion.

Parallel RCTs, when adequately powered, are generally well equipped to answer question one; however, questions two and three may not be addressed. To illustrate this, we use an example of a parallel RCT examining an intervention for dyspnea (Fig. 1A). The average intensity of dyspnea over the past 24 hours was assessed using a numeric rating scale (NRS) from 0 to 10, where 0 = no dyspnea, and 10 = worst possible dyspnea. Patients randomized to an active intervention experienced an improvement of 3 points (from 7 at baseline to 4). Patients randomized to the control intervention also reported an improvement of 2 points (from 7 at baseline to 5). Thus, the active intervention was associated with an improvement in dyspnea by one of 10. Assuming that this difference is statistically significant (question 1) because the study was adequately powered, we would then need to know if a change of one point on the NRS is clinically meaningful (question two).

Whether this difference was clinically meaningful would depend on the minimal clinically significant difference (MCID). MCID is defined as “the smallest difference in score in the domain of interest which patients perceive as beneficial and which would mandate, in the absence of troublesome side effects and excessive cost, a change in the patient’s management” (3). MCID is determined by either the anchor-based approach or the distribution based approach (4, 5). The anchor-based approach is generally preferred, and involves either asking a cohort of patients after an intervention whether their outcome of interest has changed or using an external criterion such as the frequency of rescue medication used in the case of breakthrough pain (5–7). Unfortunately, the MCID is not available for many questionnaires, making it difficult for us to know if an observed change

is clinically relevant. MCID may not always be applicable even if available because the study population and intervention often differ between the study of interest and the study in which MCID was derived. Finally, because the MCID cannot take into account individual preferences, Rennard et al. argue that it may be more appropriate for research instead of clinical practice (8).

The third measure of the effect of an intervention is overall preference. Overall preference is a pragmatic outcome because it represents a final choice, taking into account all the risks and benefits experienced by the individual, along with their relative weights. In this parallel RCT, the patients' overall preference (question three) could not be determined because patients did not have the opportunity to try both interventions.

Crossover RCTs randomize patients to a sequence of treatments, and allow investigators to overcome many of the methodological limitations of parallel RCTs to ascertain treatment differences and preferences. Crossover trials are designed for intra-individual comparisons, and participants are asked 1) to directly compare the interventions in regard to effectiveness, adverse effects and ease of use, and 2) to provide a final overall choice. Crossover trials can thus provide valuable information about which intervention is superior beyond what can be achieved in parallel studies (9). In this article, we discuss the key design elements for crossover trials, their advantages and disadvantages relative to parallel designs, and their utility in palliative care research, using a number of examples.

Advantages of Crossover Randomized Controlled Trials

In crossover RCTs, patients are randomized to receive multiple study interventions in different orders. We illustrate the design of a two-intervention, two-sequence, two-period crossover RCT in Fig. 1B. Intervention refers to the number of treatments utilized in the study, sequence to the number of different orders in which the interventions may be delivered and period to the order of the intervention within the sequence, i.e., first, second, or third. Half of the patients received the active intervention, followed by a washout period and then the control intervention, and the other half received the control intervention first, followed by a washout period and then the active intervention. Because patients have the opportunity to try all treatments, the investigators could ask them to provide their individual opinion of the changes in various scales to assess the efficacy and side effects, whether these changes are noticeable/meaningful (thus obtaining personalized MCIDs), and which treatment is preferred, taking into account the risks and benefits. Table 1 summarizes the key advantages and disadvantages of crossover trials compared to parallel studies.

In addition to providing valuable information on treatment differences and patient preferences, crossover designs offer several additional advantages, including intra-individual comparison, increased statistical power, shorter recruitment time, and cost savings.

One limitation with parallel RCTs is the wide inter-individual variations in baseline characteristics. Stratification, multivariate analysis and propensity score may allow investigators to adjust for these baseline differences to a certain extent. Crossover trials offer the distinct advantage of intra-individual comparison, thus allowing us to estimate the treatment effect more precisely while maximizing the statistical power.

In crossover RCTs, we have a greater amount of data because each patient receives multiple treatments. This, coupled with intra-individual comparison, increases the statistical power and allows a smaller number of patients compared to parallel RCTs. The following formula provides an estimate of the efficiency of crossover studies: number of patients needed in parallel study/number of patients needed in a crossover study for the same power = $2 \times (\text{between-subject variance} + \text{within-subject variance})/\text{within-subject variance}$ (10). Because within-subject variance is often less than between-subject variance, crossover studies are usually more efficient than parallel trials. In one study, the number of patients required was estimated to be 2.7- to 6.4-fold lower for a crossover study compared to a parallel design (11).

For investigators in palliative care, the smaller number of subjects in crossover RCTs is particularly attractive. First, it may take less time to enroll patients, although this may need to be balanced against the possibility of fewer interested participants because of a longer study commitment. Second, the cost of study may be lower. This is particularly important for studies in palliative care in which the clinical trials are often smaller and underfunded (12).

Finally, crossover trials are ethically sound because all participants have an equal opportunity to try all treatment interventions. In parallel trials, the investigators may sometimes adjust the ratio of randomization (i.e., 3:1 active intervention to control) so more patients have the opportunity to try the active treatment, at the cost of sacrificing statistical power.

Crossover Trial Design

In this section, we discuss some important considerations when designing crossover RCTs, with particular emphasis on the supportive/palliative care setting. Although crossover RCTs have some unique advantages, they may not be the design of choice for some research topics. When deciding whether a crossover study is appropriate for a particular study question, the investigator needs to consider several factors discussed below.

The condition being studied should be relatively stable over the study period, with the idea that the outcomes return to the baseline state during the washout period prior to starting the next intervention. In the palliative care setting, symptoms such as chronic pain, nausea, anorexia, fatigue and dyspnea are particularly suitable for crossover trials. In contrast, acute symptoms or those that fluctuate over time (e.g., delirium) may not be amendable to a crossover design.

The interventions being examined should have minimal carryover effects on the outcome of interest. For pharmacological interventions, the medications should have a relatively short half-life. Thus, crossover studies are ideal for symptom trials in which the intervention effect often resolves shortly after treatment discontinuation. Crossover studies cannot be conducted when the treatments have a permanent impact (e.g., palliative surgical procedures, antibiotics for infections), or when the intervention effect may be long lasting (e.g., counseling intervention).

Regarding the randomization sequence, a majority of the crossover trials involves only two treatments (A and B), and utilize a two-period (i.e., Period 1-Period 2) and two-sequence (i.e., sequence 1=A-B, sequence 2=B-A) design. However, the number of treatments, treatment periods and sequence permutations can vary widely depending on the study question. A crossover design is considered *uniform within sequences* if the number of times each treatment is given is the same within the sequence (e.g., A-B-C). The term *uniform within periods* is used if each treatment were given the same number of times with each period (e.g., A-B-C, B-C-A, C-A-B). Latin squares may be used to generate the sequences for crossover trials when the number of treatments and the number of periods are the same. For instance, a study consisting of three treatments [A, B, C] and three periods may use three different sequences to achieve a uniform design (e.g., A-B-C, B-C-A, C-A-B) instead of 6 different sequences ($3 \times 2 \times 1 = 6$). A crossover design is ideally balanced with respect to first-order carryover effect by ensuring that each treatment precedes the other treatment the same number of times.

The outcomes of interest should be responsive to change. Intra-individual variability needs to be similar for statistical comparison or taken into account when interpreting results. Response shift related to recalibration, reconceptualization and reprioritization of PROs may represent a potential source of bias (13, 14). Its impact may be amplified in crossover trials because of the longer study duration and repeated assessments. Further research is needed to examine this phenomenon in crossover trials. Crossover studies are not only useful for symptom trials, but also communication research and bioequivalence studies. Importantly, crossover studies cannot be conducted when the outcome of interest only occurs once (e.g., overall survival).

The washout period is engineered between study treatments to address the carryover effect from the previous treatment. During the washout period, participants return to the standard care they received prior to study initiation. This period should be long enough for patients to return to baseline before they proceed to the next intervention (15). Crossover studies with inadequate washout periods may be difficult to interpret. On the one hand, because a longer washout period can minimize the impact of any carryover effect to inadvertently affect the outcome of interest, investigators may want to lean toward the side of caution and institute a longer washout period. In pharmacological studies, five or more half-lives of the serum concentration in the blood may be used. On the other hand, some studies, such as those involving videos on communication techniques, may not require a washout period because carryover effect is generally considered negligible.

Proper blinding of patients and investigators is essential in crossover trials to minimize ascertainment bias. Appropriate masking and allocation concealment can be tested by asking the patients and investigators which treatment they felt the patient received at the end of each period. However, if a treatment has unique toxicities, this may “unmask” the trial. For crossover trials, the final treatment choice made by the patient at the end of the study after the blinded “taste test” can be particularly informative. However, some interventions cannot be easily blinded (e.g., counseling). In these cases, a final preference may still be elicited, although it may not be as objective.

The overall study duration and number of periods should take into consideration the onset and peak effect of study interventions on the condition being studied, the expected duration of the washout period, the statistical power needed and the study burden. Crossover studies, by definition, require a greater degree of commitment from participants and investigators because they will try multiple treatment options, sometimes with the same treatment more than once (e.g., four-period design). The duration of crossover studies varies widely, ranging from hours to weeks. For instance, studies on dyspnea are often short because patients are frail, the intervention effect often occurs rapidly, and the washout period is minimal. Other studies, such as those for chronic pain and nausea, may require a longer duration for the intervention to achieve optimal effect, and typically need multiple half-lives for the drugs to be metabolized and excreted during the washout period. In palliative care, patients often decline progressively. Thus, it is important to ensure that the study duration is short enough to minimize disease progression as a confounding factor. The risk of attrition also may increase with longer studies.

Statistical analysis in crossover studies is often complicated because we need to take into account the sequence effect, period effect and first-order carryover effect (10). Sequence and period effects are generally less likely to contribute to the observed differences if the study design is uniform within sequence and uniform within periods, respectively. However, their presence may have an impact on analysis and interpretation. Carryover effect can be subclassified as first-order carryover effect, and higher-order carryover effects. First-order carryover effect is defined as the residual effect from the treatment immediately preceding the current treatment, whereas higher-order carryover effects are secondary to earlier treatments before that. Higher-order effects are often assumed to be negligible in statistical models. Collaboration with a biostatistician during the design phase is essential to ensure proper sample size calculation and a comprehensive analysis plan.

Data from crossover studies can be analyzed at three levels: group comparison, individual responder analysis, and patient preference. In group comparison, the classical Grizzle analysis involves the two-sample *t*-tests (or Wilcoxon rank-sum test) to evaluate the treatment effect, period effect, and interaction effect (16). This is equivalent to an analysis of variance, with independent variables being patient, treatment and time. There may be a need to convert the analysis using data only from the first period as if it were a parallel RCT if the attrition rate is high or the carryover effect is prolonged (17). Alternatively, linear mixed models may be used to examine variables that are associated with the outcome, with the factor of interest being the study treatment, fixed effects being the period effect and the time by treatment interaction, and random effects being individual patients across time where correlation within patient is modeled using an appropriate covariance structure. In individual responder analysis, we can determine the intra-individual treatment effect and the proportion of responders based on predefined cutoffs for each study arm. Paired *t*-test may be used to determine the statistical significance of within-arm changes. Finally, the preference of patients after having tried all study interventions can be highly informative and pragmatic.

A number of variations in crossover study designs merit discussion. N-of-1 trials are crossover designs that involve a single subject randomized to receive the study interventions in different orders (e.g., A-B-B-A-B-A-A-B) (18–21). Sometimes, multiple patients may

each complete an N-of-1 trial, and the data is reported in aggregate (22). Unlike traditional crossover designs, N-of-1 trials often have at least four periods, the unit of randomization is the treatment order instead of the patient, and the emphasis is within-subject comparison instead of group comparison (23). N-of-1 trials may be particularly useful to examine rare symptoms in palliative care populations (24, 25). Wait-list control trials, or delayed “crossover” studies, employ an observation period as the control. Thus, patients are assigned to either early treatment group (i.e., treatment intervention with or without an observation period afterwards) or late treatment group (an observation period followed by study intervention). These designs are often used to examine complex health interventions that cannot be blinded to patients, have a short duration, and limited availability (26–28). Unlike traditional crossover studies, wait-list control studies are usually analyzed as parallel trials or within-arm comparisons. A recent meta-analysis of psychotherapy trials suggested that a wait list control may exaggerate treatment response by introducing a placebo effect, which warrants further investigation (29).

Examples of Crossover Trials in Supportive/Palliative Care Research

In the following section, we use examples of several crossover RCTs in supportive/palliative care conducted by our group to illustrate the key design elements, pearls and pitfalls related to study design and analysis for these trials.

Patient Preference vs. Treatment Benefits

In a double-blind, randomized crossover study, Bruera et al. (30) compared the efficacy of mazindol and placebo on various symptoms. Among 26 evaluable patients, mazindol was associated with improved pain, worsened anxiety and appetite, and similar levels of activity and depression. At the end of the study, 10 (38%) patients chose mazindol and seven (27%) chose placebo; nine (35%) investigators selected mazindol and 11 (42%) selected placebo.

This study shows how patient preference can provide information beyond a quantitative comparison of efficacy. Although pain control was better, the other symptoms were worse. Overall, the number of patients favoring mazindol and placebo were similar. Based on these findings, mazindol was not recommended as a treatment for pain.

Patient Preference

Bruera et al. (31) conducted a four-period, double-blind, crossover RCT comparing the effect of oxygen 5 L/min for five minutes versus air 5 L/min for five minutes on dyspnea in hypoxemic cancer patients. Fourteen patients were randomized to one of four arms: (A) oxygen → air → oxygen → air; (B) oxygen → air → air → oxygen; (C) air → oxygen → oxygen → air; and (D) air → oxygen → air → oxygen. The primary outcome measure was a dyspnea visual analogue scale that ranges from 0 (no dyspnea) to 100 (worst dyspnea). After each pair of oxygen-air treatments, patients and investigators were asked to choose which intervention was more effective.

Oxygen was found to be more effective than air in alleviating dyspnea by 20.5 mm on a visual analogue scale ($P<0.001$). Importantly, 12 of 14 patients chose oxygen consistently

for both paired comparisons. In 12 of 14 cases, investigators also consistently picked oxygen in both paired comparisons.

This study highlights the use of a four-period trial to achieve a high statistical power, and the use of patient preference to provide valuable information on treatment efficacy.

Washout Period

In a feasibility crossover RCT (32), hospitalized patients with advanced cancer and refractory dyspnea were randomized in a 1:1 ratio to receive either (A) two hours of high flow oxygen (HFOx) followed by a variable washout period and then two hours of bilevel positive airway pressure (BiPAP), or (B) two hours of BiPAP followed by a variable washout period and then two hours of HFOx. This study was not blinded because of the nature of study interventions.

As part of the study design, we instituted a variable washout period after the first intervention to determine the optimal duration required for patients to return to their baseline dyspnea level. This was conducted by checking the intensity of dyspnea every 10 minutes. Patients were able to proceed to the second intervention if and when their dyspnea level was lower or equal to the baseline level of intensity + 1. We limited the maximum duration of the variable washout period to one hour because previous studies on dyspnea of supplemental oxygen revealed that patients return to baseline in a relatively short time frame.

Thirty patients were enrolled. Among the patients who completed the first intervention and washout period, two of 10 patients who received HFO and four of seven patients who received BiPAP reported significant improvement of dyspnea for greater than one hour, making them ineligible to proceed to the second intervention. Thus, the study was amended to be analyzed as a parallel study. Both HFOx and BiPAP were associated with significant improvements in dyspnea NRS scores, with a mean change of 2.1 and 3.2 clinically.

The observation of a prolonged effect on dyspnea well after completion of the intervention is of interest, and warrants further investigation. This study highlights the importance of having an adequate washout period, the need to analyze a study as parallel if a large portion of patients did not continue onto the second intervention because of longer than expected therapeutic effect or dropouts.

Period Effect

We investigated the effect of physician posture when breaking bad news on patients' perception of their compassion (33), using a crossover RCT design. One hundred seventy-three cancer patients were randomly assigned to watch two nine-minute videos in different order. The two videos were identical except for physician posture; one involved the physician sitting, and in the other, the physician was standing. At the end of the study, patients were asked to choose which session they preferred. The research nurses were blinded to the allocation sequence throughout the study and patients were blinded to the study hypothesis. There was no washout period.

The perceived level of compassion was higher for the sitting physician compared to the standing physician (33 vs. 29, $P < 0.0001$). Importantly, 85 (51%) patients preferred the sitting physician, whereas 29 (17%) preferred the standing physician. We also found significant period effect, in which the second consultation was associated with higher ratings of compassion (32 vs. 30, $P = 0.003$) and greater overall preference (odds ratio 3.2–3.4). This period effect also was observed in a similar study (34).

Crossover designs are not limited to symptom trials, but can be helpful in communication research (35). Importantly, this study illustrates that period effect needs to be taken into consideration in the analysis, and could have important clinical implications.

Summary

By randomizing patients to a sequence of treatments, crossover studies offer many potential advantages over parallel designs, including the ability to ascertain clinical significance and elicit patient preferences, intra-individual comparisons, increased statistical power, faster completion, lower cost, and equality for trial participants. However, crossover trials are only feasible when the condition being studied is relatively stable and the intervention effect is short term with limited carryover effect. The longer study duration may also result in suboptimal rates of recruitment and retention. Crossover studies with inadequate washout periods may be difficult to interpret. Balancing the advantages and disadvantages, crossover designs are appropriate in supportive/palliative care to determine pharmacokinetics (bioequivalence studies) and to assess the efficacy of various interventions in symptom and communication research.

Acknowledgments

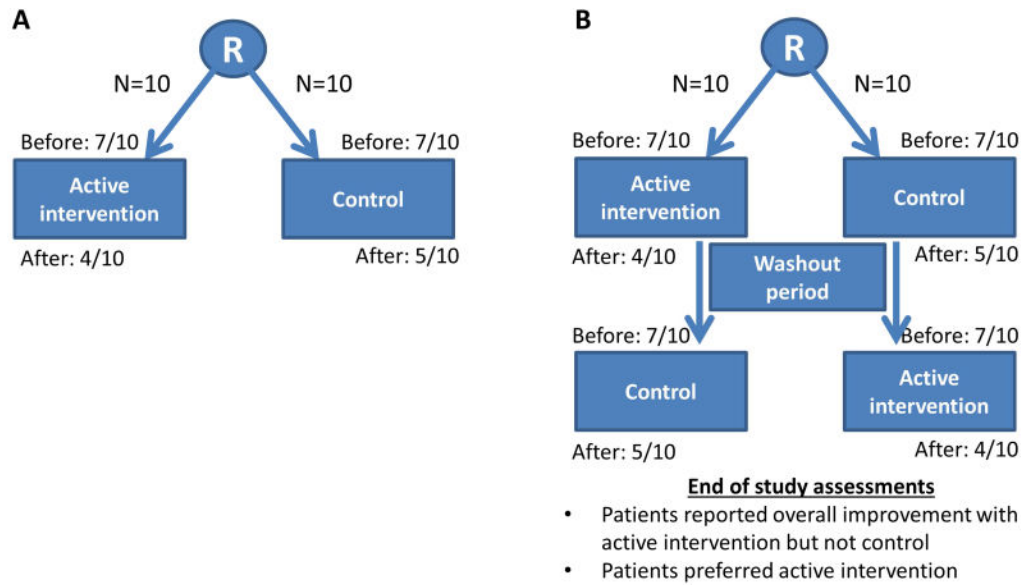
This work was supported in part by National Institutes of Health grants RO1NR010162-01A1, RO1CA122292-01, and RO1CA124481-01 (Dr. Bruera). This study was also supported by the M. D. Anderson Cancer Center Support Grant (CA 016672) and an institutional startup grant #18075582 (Dr. Hui).

References

1. Hui D, Parsons HA, Damani S, et al. Quantity, design, and scope of the palliative oncology literature. *Oncologist*. 2011; 16:694–703. [PubMed: 21471275]
2. Hui D, Arthur J, Dalal S, Bruera E. Quality of the supportive and palliative oncology literature: a focused analysis on randomized controlled trials. *Support Care Cancer*. 2012; 20:1779–1785. [PubMed: 21935717]
3. Jaeschke R, Singer J, Guyatt GH. Measurement of health status. ascertaining the minimal clinically important difference. *Control Clin Trials*. 1989; 10:407–415. [PubMed: 2691207]
4. Wells G, Beaton D, Shea B, et al. Minimal clinically important differences: review of methods. *J Rheumatol*. 2001; 28:406–412. [PubMed: 11246688]
5. Copay AG, Subach BR, Glassman SD, Polly DW Jr, Schuler TC. Understanding the minimum clinically important difference: a review of concepts and methods. *Spine J*. 2007; 7:541–546. [PubMed: 17448732]
6. Redelmeier DA, Guyatt GH, Goldstein RS. Assessing the minimal important difference in symptoms: a comparison of two techniques. *J Clin Epidemiol*. 1996; 49:1215–1219. [PubMed: 8892486]
7. Hui D, Bruera E. Minimal clinically important differences in the Edmonton Symptom Assessment System: the anchor is key. *J Pain Symptom Manage*. 2013; 45:e4–5.

8. Rennard SI. Minimal clinically important difference, clinical perspective: an opinion. *COPD*. 2005; 2:51–55. [PubMed: 17136962]
9. Mazzocato C, Sweeney C, Bruera E. Clinical research in palliative care: choice of trial design. *Palliat Med*. 2001; 15:261–264. [PubMed: 11407200]
10. Wellek S, Blettner M. On the proper use of the crossover design in clinical trials: part 18 of a series on evaluation of scientific publications. *Dtsch Arztebl Int*. 2012; 109:276–281. [PubMed: 22567063]
11. Margarone JE, Liebow C, Hall RE, Braun RE, Weiner M. A comparison of crossover versus parallel-group design in the evaluation of analgesic efficacy after molar extraction. *Clin Pharmacol Ther*. 1995; 58:453–458. [PubMed: 7586938]
12. Hui D, Reddy A, Parsons HA, Bruera E. Reporting of funding sources and conflict of interest in the supportive and palliative oncology literature. *J Pain Symptom Manage*. 2012; 44:421–430. [PubMed: 22771126]
13. Hagedoorn M, Sneeuw KC, Aaronson NK. Changes in physical functioning and quality of life in patients with cancer: response shift and relative evaluation of one's condition. *J Clin Epidemiol*. 2002; 55:176–183. [PubMed: 11809356]
14. Howard JS, Mattacola CG, Howell DM, Lattermann C. Response shift theory: an application for health-related quality of life in rehabilitation research and practice. *J Allied Health*. 2011; 40:31–38. [PubMed: 21399850]
15. Matthews JN. Multi-period crossover trials. *Stat Methods Med Res*. 1994; 3:383–405. [PubMed: 7889228]
16. Grizzle JE. The two-period change-over design and its use in clinical trials. *Biometrics*. 1965; 2:469–480.
17. Willan AR, Pater JL. Carryover and the two-period crossover clinical trial. *Biometrics*. 1986; 42:593–599. [PubMed: 3567292]
18. O'Brien PC. The use and misuse of N-of-one studies. *Int J Oral Maxillofac Implants*. 1997; 12:293. [PubMed: 9197093]
19. Vohra S, Punja S. N-of-1 trials: individualized medication effectiveness tests. *Virtual Mentor*. 2013; 15:42–45. [PubMed: 23356806]
20. Duan N, Kravitz RL, Schmid CH. Single-patient (n-of-1) trials: a pragmatic clinical decision methodology for patient-centered comparative effectiveness research. *J Clin Epidemiol*. 2013; 66:S21–28. [PubMed: 23849149]
21. Gabler NB, Duan N, Vohra S, Kravitz RL. N-of-1 trials in the medical literature: a systematic review. *Med Care*. 2011; 49:761–768. [PubMed: 21478771]
22. Cushing CC, Walters RW, Hoffman L. Aggregated N-of-1 randomized controlled trials: modern data analytics applied to a clinically valid method of intervention effectiveness. *J Pediatr Psychol*. 2014; 39:138–150. [PubMed: 24284134]
23. Sedgwick P. What is an “n-of-1” trial? *Br Med J*. 2014; 348:g2674.
24. Bruera E, Schoeller T, MacEachern T. Symptomatic benefit of supplemental oxygen in hypoxemic patients with terminal cancer: the use of the N-of-1 randomized controlled trial. *J Pain Symptom Manage*. 1992; 7:365–368. [PubMed: 1517652]
25. Nikles J, Mitchell G, Walters J, et al. Prioritising drugs for single patient (n-of-1) trials in palliative care. *Palliat Med*. 2009; 23:623–634. [PubMed: 19605605]
26. Farquhar M, Higginson IJ, Booth S. Fast-track trials in palliative care: an alternative randomized controlled trial design. *J Palliat Med*. 2009; 12:213. [PubMed: 19254192]
27. Siddons HM, Wootten AC, Costello AJ. A randomised, wait-list controlled trial: evaluation of a cognitive-behavioural group intervention on psycho-sexual adjustment for men with localised prostate cancer. *Psychooncology*. 2013 Apr 10. [Epub ahead of print].
28. Zernicke KA, Campbell TS, Specia M, et al. A randomized wait-list controlled trial of feasibility and efficacy of an online mindfulness-based cancer recovery program: the eTherapy for cancer applying mindfulness trial. *Psychosom Med*. 2014; 76:257–267. [PubMed: 24804884]
29. Furukawa TA, Noma H, Caldwell DM, et al. Waiting list may be a nocebo condition in psychotherapy trials: a contribution from network meta-analysis. *Acta Psychiatr Scand*. 2014; 130:181–192. [PubMed: 24697518]

30. Bruera E, Carraro S, Roca E, Barugel M, Chacon R. Double-blind evaluation of the effects of mazindol on pain, depression, anxiety, appetite, and activity in terminal cancer patients. *Cancer Treat Rep.* 1986; 70:295–298. [PubMed: 3512080]
31. Bruera E, de Stoutz N, Velasco-Leiva A, Schoeller T, Hanson J. Effects of oxygen on dyspnoea in hypoxaemic terminal-cancer patients. *Lancet.* 1993; 342:13–14. [PubMed: 8100289]
32. Hui D, Morgado M, Chisholm G, et al. High-flow oxygen and bilevel positive airway pressure for persistent dyspnea in patients with advanced cancer: a phase II randomized trial. *J Pain Symptom Manage.* 2013; 46:463–473. [PubMed: 23739633]
33. Bruera E, Palmer JL, Pace E, et al. A randomized, controlled trial of physician postures when breaking bad news to cancer patients. *Palliat Med.* 2007; 21:501–505. [PubMed: 17846090]
34. Strasser F, Palmer JL, Willey J, et al. Impact of physician sitting versus standing during inpatient oncology consultations: patients' preference and perception of compassion and duration. A randomized controlled trial. *J Pain Symptom Manage.* 2005; 29:489–497. [PubMed: 15904751]
35. Rhondali W, Perez-Cruz P, Hui D, et al. Patient-physician communication about code status preferences: a randomized controlled trial. *Cancer.* 2013; 119:2067–2073. [PubMed: 23564395]



Interpretation of results

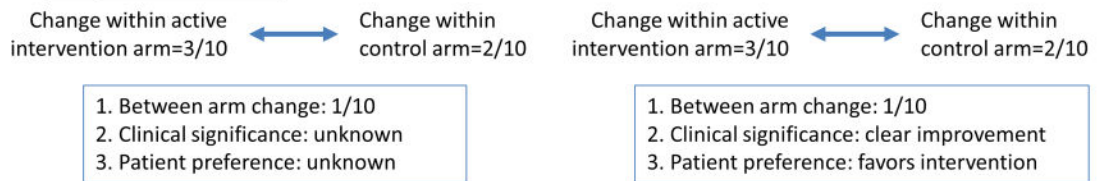


Fig. 1. Randomized controlled trial design. (A) Parallel trial, (B) Crossover trial.

Table 1

Advantages and Disadvantages of Crossover Randomized Trials Relative to Parallel Trials

Advantages	Disadvantages
<ul style="list-style-type: none"> • Inter-individual group comparisons (reduced confounding factors) • Intra-individual responder analysis possible • Overall choice/preference can be obtained • All patients have the opportunity to try all study interventions • Smaller sample size to achieve the same statistical power—potentially shorter recruitment period and less costly trial 	<ul style="list-style-type: none"> • Limitation in condition being studied (i.e., should be relatively stable over time) • Limitation in interventions being studied (i.e., should have limited carryover effect and no permanent effect on the outcome) • Limitation in outcome measures (i.e., should be responsive to change and cannot occur only once) • Prolonged trial duration (multiple interventions and washout periods) • Response shift may be amplified • Potential for higher attrition rates, which could significantly diminish the power of the study

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript