Published in final edited form as:
Planta Med. 2011 September ; 77(13): 1473–1476. doi:10.1055/s-0030-1271044.

Control in the Middle (CIM) for Three Period Crossover Studies

Jonathan Shuster[1], **Stephen D. Anton**[2,3], **Douglas Theriaque**[4], and **Saunjoo Yoon**[5]

[1]Department of Health Outcomes and Policy, College of Medicine (COM), University of Florida, Gainesville, Florida, USA

[2]Department of Aging and Geriatric Research (COM), University of Florida, Gainesville, Florida, USA

[3]Department of Clinical and Health Psychology, University of Florida, Gainesville, Florida, USA

[4]Clinical and Translational Science Institute (COM), University of Florida, Gainesville, Florida, USA

[5]Department of Adult and Elderly, College of Nursing, University of Florida, Gainesville, Florida, USA

Abstract

Three period crossover studies can be efficient and convenient methods of conducting Phase II clinical trials. Non-randomly placing control in the middle (CIM) has not been practiced, but may be extremely useful in studies testing herbal products for which placebos are not available, or for distinguishing between behavioral and biological effects. Furthermore, this design can serve as a valuable addition to classical studies of either (a) two competing treatments or (b) treatment versus placebo versus an open label “nothing” as the control. Therefore, we propose rigorous designs that will help practitioners efficiently answer research questions where (1) two active treatments need to be compared against each other with treatment vs. placebo comparisons of secondary importance; (2) a single active treatment needs to be tested where no placebo is available; or (3) the placebo effect is of interest in a treatment vs. placebo trial. For studies where no placebo is available, deception will be required, with participants told that in one randomly selected period (#1 or #3) they will receive the active treatment, and that they will receive a new experimental inert placebo in the other period. Assuming this design is approved by an ethics committee, it can be very useful in biomedical research.

Keywords

blinding; crossover study; deception; phase II clinical trial; placebo; washout

In this paper, we propose a novel design concept, namely non-randomly placing control in the middle (CIM), for three period crossover designs with two treatments plus a control. One reason to place control in the middle is to unify the washout time between the two active treatments. A second reason, which might be perceived as controversial, is that this design could be useful when there is no placebo available for an active treatment. Deception would be needed, where the subject gets the same experimental treatment in Periods #1 and #3, nothing in Period #2, but is told on one randomly selected period either (R) s/he is getting the real treatment or (P) s/he is getting a new inactive placebo being tested for the new

Correspondence Jonathan J. Shuster, PO Box 100177, University of Florida, Gainesville, FL 32610-0177, USA. jshuster@biostat.ufl.edu. Phone: (352) 265-0111 Ext. 86503.

Conflict of Interest: None of the authors have any conflicts of interests regarding this article.

treatment. A third reason, perhaps the most compelling, is for classical active treatment (R) versus Placebo (P) crossover studies, whereby extending the washout interval slightly would allow for the assessment of the "placebo effect." These designs will be discussed in detail in the next section, followed by recommended methods for data analysis and by a discussion.

RATIONALE FOR CONTROL IN THE MIDDLE (CIM) DESIGNS

There are a number of considerations for a crossover study design involving two active treatments plus placebo(s), where all three treatment contrasts are of interest. In a recent grant application, we proposed a rather standard randomized block design assigning one sixth of the participants to each of the six orderings (permutations) of A+Placebo B, B +Placebo A, or Placebo A + Placebo B. A "washout" interval of eight weeks was inserted between each period as an attempt to avert carryover effects. See Figure 1 for details. The National Center for Complementary and Alternative Medicine (NCCAM) oversight committee had two concerns about this design. First, because there was a possibility that the proposed agents have long acting pharmacological effects, it was suggested that an eight week washout after active treatment may not be sufficient to eliminate or at least minimize residual effects. Another potential competing concern with this design was that the total study time (31 weeks consisting of three five-week treatment periods separated by two eight-week washout intervals) was somewhat longer than ideal, and could lead to higher attrition rates than desirable. One potential solution, suggested to address both of these concerns, was to split the single study crossover design into two, two-period crossover studies, one for each active treatment vs. placebo. Participants would be randomly assigned in equal numbers to one of these two crossover designs. This design is depicted in Figure 2. This would allow for comparison of the effects of the active treatments in a randomized parallel study. Unfortunately, the loss of the repeated measures aspect of the study, especially given that the comparison of the active treatments was the most important contrast, made the sample size requirements for this approach well beyond our accrual capability.

A third option, depicted in Figure 3, is a randomized crossover study design of the active treatments in periods #1 and #3, with placebo(s) non-randomly assigned in the middle period. This would allow for a uniformly timed washout interval between the two active treatments. It also preserves participants' blinding for all periods, and investigator blinding of the active treatments. In this design, participants are randomly assigned to one of the active treatments during an initial five week period, to a four week washout interval, to five weeks of placebo, to a second four week washout interval, and finally to the other active treatment during the final five week time period. This design reduces the total study time by eight weeks (total study duration is 23 weeks) and unifies the true washout time between active treatments from eight weeks for two thirds of the participants and 16 weeks for the other third to 13 weeks for all participants, making it responsive to potential concerns that eight weeks was not an adequate washout interval. This design also preserves a true randomized crossover design for each contrast, double blinding of the active treatments, and single blinding of the placebo contrasts.

Complementary and Alternative Medicine trials

In some potential studies of herbal products, there may be no placebo available. Unlike pharmaceuticals, these products can only be marketed as dietary supplements in the United States (Dietary Supplement Health and Education Act, 1994), not as a New Drug Indication (IND), which invariably requires a placebo formulation. In addition, these products are widely available in health food stores and via the internet, and have very distinctive tastes and odors that may be recognized by the research subject, or even investigated over the internet by the research subject. In such studies, especially with soft endpoints (Likert

Scales, for example), it is vital to maintain maximal blinding. This can be accomplished effectively by a CIM design, although some subject deception is required. Therefore, this design needs to be limited to products known to be safe in the dosage and duration being studied.

The participants will be given the same experimental treatment in the identical dosage in Periods #1 and #3. Participants will get nothing in Period#2. The participants will be randomly told before starting Period #1, either (R) You are getting the real treatment or (P) you are getting an inert experimental placebo for the treatment. They will be crossed over just before Period #3, as to what is told to them.

The major comparison would pit Period #2 vs. (P), as a crossover analysis. Secondarily, contrasting (R) vs. (P) provides information on the "Placebo Effect."

Two-period Crossover Designs Adapted to CIM

In the classical two period crossover design, either an active treatment against a placebo or two active treatments (including two doses of the same active treatment), by extending the washout interval, one can convert this to a CIM design. For placebo controlled studies, the middle period can be compared to the placebo period to assess the placebo effect. For active control studies, the middle period might be placebo(s) to enable assessment of the effect of each active treatment without adding significant expense to the study.

Statistical Methods for CIM Designs

Since the research questions all boil down to two-treatment comparisons, standard analytic methods for the design and analysis of two-period crossover designs can be utilized - See Senn [1], Shuster [2], and Tudor and Koch [3] for further details. For reasons of robustness against carryover effects, period effects, and dropout rates, as well as for simplicity, we recommend the method of Shuster [2]. If the subscripts "Late" and "Early" represent the result from the later period and earlier period for the two specific treatments being compared, and j represents subject #j, we define the dependent variable $D_j=0.5(Y_{Late,j}-Y_{Early,j})$, with Y the value of the dependent variable.

The two orderings (AB vs. BA for example) are compared by the two-sample T-test for quantitative variables, the two sample Wilcoxon test for ordinal data, or by two sample Z-test for dichotomous data (e.g., Yes/No).

The difference in the sample means of the D_j's for the two orderings, absent missing data, is an unbiased estimate of the unweighted average of the effect sizes in the early and late periods, irrespective of whether the randomization turned out to be 50–50, or whether there were period or carryover effects. Furthermore, as shown in Shuster [4], the t-test and Wilcoxon test yield large sample approximations to their respective permutation tests of the strong null hypothesis that the distribution of D_j is the same for both orderings. For sufficient sample sizes this makes them assumption-free.

Note that under an even stronger null hypothesis that the participant's fate is predestined (i.e. would have been the same irrespective of the ordering), then the T-test, Wilcoxon Test, and Z-test still are valid approximations to permutation tests, with participants with missing values for one or both of the treatments being compared, legitimately dropped from the analysis.

Since the hypotheses are tested by standard two-sample tests, classical power analysis can be applied to these studies. However, if a user has additional information on the planning distribution under the null and alternative hypotheses, the methods of Shuster [5] would be

very useful to contrast the actual power for the t-test vs. the Wilcoxon test to lock in a sample size and method before beginning the study. SAS macros to accomplish this are available at http://ags.bwh.harvard.edu/, under the button “Jon Shuster’s SAS design and analysis programs.”

DISCUSSION

Although to the best of our knowledge, no study expressly using a CIM design has ever been published, a paper published by Chow and Shao [6] was brought to our attention by the editor, where a three period crossover design was advocated with Period #2 non-randomly repeated as Period #3. Lacking a published precedent, we present two published randomized crossover studies of complementary and alternative medicine that might have been enriched with the CIM design. Kuratsune and colleagues [7] conducted a placebo controlled double blind two period crossover study of the effect of crocetin from Gardenia Jasminoides Ellis upon episodes of wakening (per actigraphy) and subjective sleep symptoms amongst sleep deprived participants. The treatments were assessed over two weeks, with a two week washout. Had the washout been extended to four weeks, with an assessment during Weeks 4 and 5, the middle two weeks of the washout, the investigators would have been able to assess the placebo effect (placebo vs. open label nothing) and the overall effect of the active treatment (treatment vs. nothing). The extra two weeks makes the design symmetric in terms of having a one week washout period of nothing before and after the observation period of two weeks of the open label period of nothing. If, for example, the dependent variable is change from the period’s baseline, one would prefer the middle period to be from a real no agent baseline, and not from the end of a real agent (for half) or placebo (for half).

In the second study, Abdul and colleagues [8] conducted an open label randomized three period crossover study of warfarin alone vs. warfarin plus cranberry vs. warfarin plus garlic. The major research question was whether either of these substances interacted with warfarin pharmacodynamically. The design was efficient in that they were able to utilize the same control for the impact of two food substances. For the period where the subject was on warfarin plus an active treatment, the subject received a two week period of the complimentary product alone followed by a week on the product plus warfarin. For the period on warfarin alone, there was a one week treatment. There was a two week washout interval between treatments periods. The primary statistical comparison was to test the impact of each treatment on the warfarin pharmacodynamics, and only secondarily to compare the impact of cranberry vs. garlic. Despite its efficiency, the study had some asymmetry induced by its classical design. First, if warfarin alone was given in the second or third period, it would follow the previous period’s warfarin by two weeks, whereas if warfarin alone was administered in the first or second period, the next warfarin would be given four weeks later. Further, the time separation for any pair of treatments was not internally consistent, being five weeks apart or ten weeks apart, depending upon the randomization order. Now consider a CIM design where warfarin alone was placed non-randomly in the middle, including a two week period of nothing prior to warfarin alone. Preferably, if placebos were available for each product, a double placebo period of two weeks is introduced prior to warfarin alone, with single placebo added to each active treatment. Now the timing is entirely consistent both between evaluations and within evaluations. For the major comparisons, there is no confounding influence of the garlic period for some participants being between the warfarin alone period and the cranberry period (and vice versa).

Two other studies provide strong motivation to the reality of a placebo effect, something that CIM designs could be very useful to ferret out. Kaptchuk and colleagues [9] looked at what was effectively a parallel randomized study of run in periods on a sham device and on a

placebo pill and found a striking difference in the trajectory of subjective pain for the two groups, favoring the sham device over the placebo pill. It would in fact be of interest to run a replication of this study via a CIM design with the control open label nothing. In the second article, Mayberg and colleagues [10] compared fluoxetinean vs. placebo for unipolar depression. They found meaningful but similar improvement in depression symptoms as measured by objective means for both groups. Since this was a small study of relatively short duration, it would be of interest to verify these conclusions in a CIM study, which includes open label nothing between randomized ordering of the placebo and the active treatment.

The CIM design, with placebo or nothing in the middle, provides an excellent diagnostic tool for future research in the study's subject area, as it has three evaluations absent active treatment, that follow an active treatment (after the first washout interval, after the placebo/nothing period, and after the second washout interval). This allows assessment of the question as to response being associated with time since cessation of the active treatment. Another good feature of placing placebo in the second period is the enhanced ability to compare side effects between the active treatments. Although the treatment may wash out quickly, full recovery from side effects may take longer, inducing a greater potential for carryover side effects in the design in a pure three period crossover design than the CIM design. This is especially important when the active treatments represent two doses of a particular treatment, a common requirement in Complementary and Alternative Medicine trials.

If enough of these studies can be conducted, we can subject them to a random effects meta-analysis to determine an overall assessment of the average placebo effect.

There are drawbacks to CIM designs as compared to standard two period crossover designs with a washout. First, extending the duration of the study observation time will increase the likelihood of dropouts. Second, the added complexity may adversely affect accrual. Third, the middle period is generally single blind (participants are blinded but not researchers), even if the other periods are double blind. There is one additional price to be paid by the CIM design with placebo in the middle. There is potential confounding by the fact that unlike participants evaluated for the Period #1 to Period #2 differences, participants evaluated for the Period #2 to Period #3 differences have previously been exposed to the other active treatment. This issue also exists for classical three period designs, but there is a balance that this occurs approximately equally often in each direction. These issues should be considered before one adopts a CIM design.

Although it is not within the scope of this paper, researchers that implement the CIM design, should employ all possible measures to minimize dropouts. Statistically, dropouts can be accommodated in the analysis per the previous section under the strong null hypothesis stated, but if dropout is treatment related, then some bias in effect size point and interval estimates could occur. Irrespective of compliance, every effort needs to be made to obtain the data at all three observation points, so that intent-to-treat analysis can be approximated to the best extent possible. As for the single blind issue for the middle period (participants but not researchers blinded), it may actually be possible to blind the investigators in the middle period, but that will create a second level of deception, where the biostatistician, ethics committee, and investigational pharmacy would work together to produce a protocol written as if the study was a randomized block design (participants see all three treatments in random order), while in practice, it is a CIM design.

The CIM design with sham placebo does require deception because participants are told they are guaranteed to receive the "placebo" in one of the periods. However, this could be

approved by an Ethics Committee if they are convinced that (1) giving the active drug for two periods keeps the trial in the “No more than minimal risk” category and (2) scientific justification for the use of the CIM design over a conventional design is acceptable to the committee.

We conclude that the CIM design can be a very effective tool, especially if one of three design scenarios occurs: (1) the major goal is to compare the two active treatments, including two doses of the same treatment; (2) there is no placebo available for the product; or (3) the researcher has an interest in evaluating the placebo effect. These studies can be analyzed in a straight forward manner. If deception is involved, in order to avert any operative issues, we recommend contact with the ethics committee on this issue be completed prior to writing a protocol.

Abbreviations

CIM Control in the middle design

Acknowledgments

This work was partly supported by National Center for Research Resources, National Institutes of Health; grant numbers: M01RR00082, 1UL 1RR029890, U54RR025208, K23AT004251, American Heart Association (09CRP2390173), and Thomas H. Maren Foundation. Finally, we would like to extend our appreciation to Pam Selby, Editor, University of Florida College of Nursing, for her excellent job of internal editing the manuscript.

References

1. Senn, SJ. Cross-over Trials in Clinical Research. Chichester: Wiley; 2002.
2. Shuster, JJ. Ambrosius, WT. Chapter 12 in Methods in Molecular Biology: Elementary Biostatistics. Totawa: Humana Press; 2007. Design of Experiments.
3. Tudor G, Koch GG. Review of nonparametric methods for the analysis of crossover studies. Stat Methods Med Res. 1994; 3:345–381. [PubMed: 7889227]
4. Shuster JJ. Diagnostics for assumptions in moderate to large simple clinical trials: Do they really help? Stat Med. 2005; 24:2431–2438. [PubMed: 15977289]
5. Shuster JJ. Student T-tests for potentially abnormal data. Stat Med. 2009; 28:2170–2184. [PubMed: 19326398]
6. Chow S, Shao J. Statistical methods for two-sequence three-period crossover designs with incomplete data. Stat Med. 1997; 16:1031–1039. [PubMed: 9160497]
7. Kuratsune H, Umigai N, Takeno R, Kajimoto Y, Nakano T. Effect of crocetin from Gardenia jasminoides Ellis on sleep: a pilot study. Phytomedicine. 2010; 17:840–843. [PubMed: 20537515]
8. Abdul MIM, Jiang X, Williams KM, Day RO, Roufgalis WSBD, Liauw HX, McLachlan AJ. Pharmacodynamic interaction of warfarin with cranberry but not with garlic in healthy subjects. Br J Pharmacol. 2008; 154:1691–1700. [PubMed: 18516070]
9. Kaptchuk TJ, Stason WB, Davis RB, Legedza ATR, Schnyer RN, Kerr CE, Stone DA, Nam BH, Kirsch I, Goldman RH. Sham device *vs*. inert pill: randomised controlled trial of two placebo Treatments. Br Med J. 2006; 332:391–397. [PubMed: 16452103]
10. Mayberg HS, Silva JA, Brannan SK, Tekell JL, Mahurin RK, McGinnis S, Jerabek PA. The Functional Neuroanatomy of the Placebo Effect. Am J Psychiatry. 2002; 159:728–737. [PubMed: 11986125]

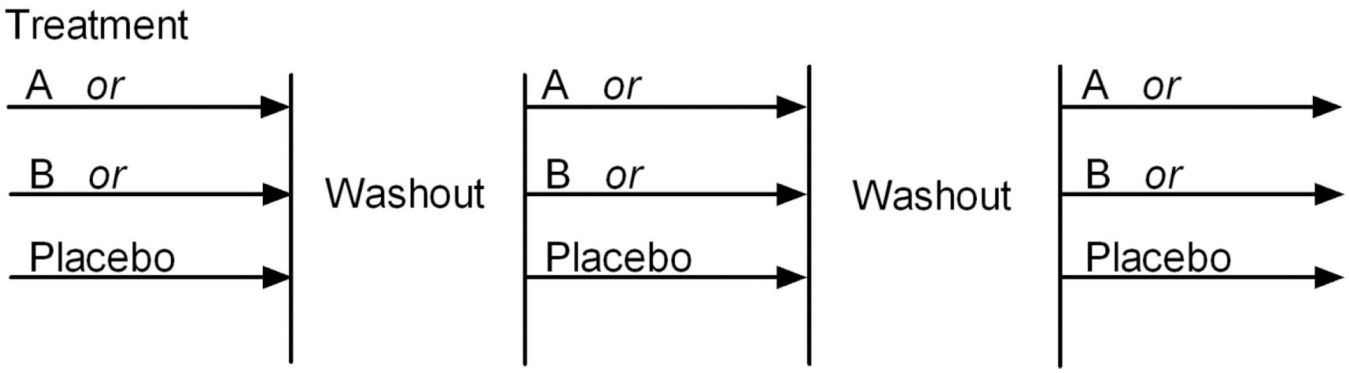

Figure 1. Randomized Block Design

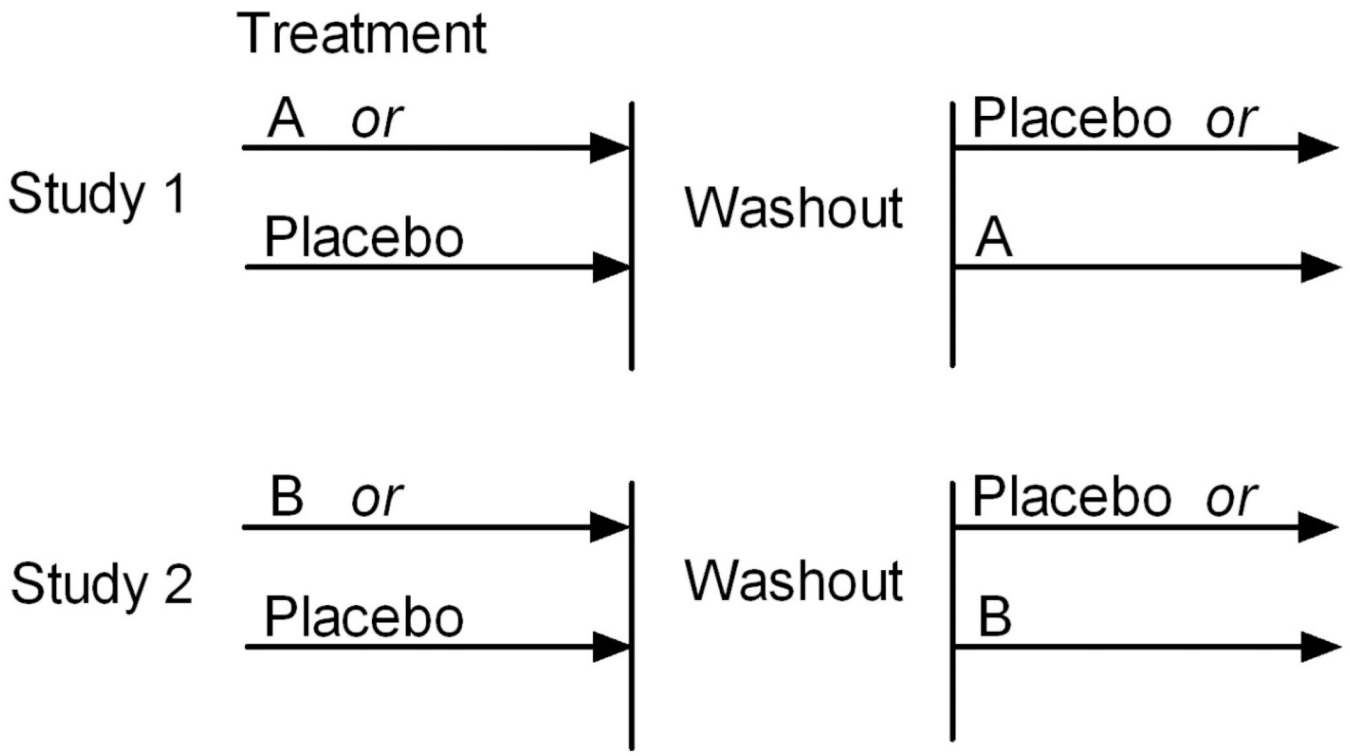

Figure 2. Two Two-Period Crossover Designs

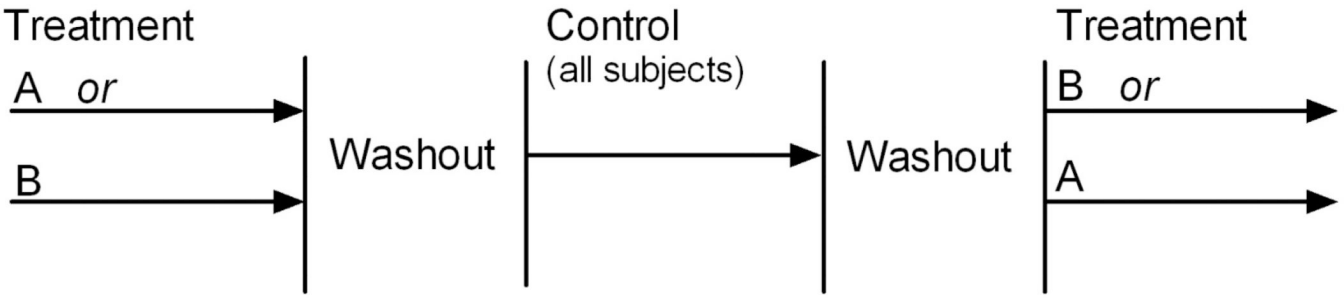

Figure 3. Control-in-the-Middle Design