Title
The misguided ethics of crossover trials

Journal
Contemporary Clinical Trials, 37(2)

ISSN
1551-7144

Authors
Prasad, Vinay
Grady, Christine

Publication Date
2014-03-01

DOI
10.1016/j.cct.2013.12.003

Peer reviewed

Published in final edited form as:
Contemp Clin Trials. 2014 March ; 37(2): 167–169. doi:10.1016/j.cct.2013.12.003.

The Misguided Ethics of Crossover Trials

Vinay Prasad, MD and
Medical Oncology Branch, National Cancer Institute, National Institutes of Health, 10 Center Dr. 10/12N226, Bethesda, MD 20892, Phone: 219.229.0170

Christine Grady, RN, PhD
Department of Bioethics, National Institutes of Health, 10 Center Dr. 10/1C118, Bethesda MD 20892, Phone 301.496.2429

Vinay Prasad: vinayak.prasad@nih.gov; Christine Grady: cgrady@cc.nih.gov

Abstract

Crossover is increasingly favored in trials of cancer therapies; even those that seek to establish the basic efficacy of novel drugs. Crossover is done in part for trial recruitment, but also out of a sense of doing the right thing—offering the investigational agent to more patients. In this paper, we argue that this ethical feeling—that crossover is a preferred trial choice—is misguided. In seeking to sate the desires of participants, we might undermine a trial's ability to answer a meaningful clinical question. When a trial is incapable of answering a question, it becomes unethical. Using a crossover strategy in oncology clinical trials can make trials less ethical, not more.

L'enfer est plein de bonnes volontés et désirs

(Hell is full of good wishes and desires)

--Saint Bernard of Clairvaux (c.1150)

An estimated quarter of randomized controlled trials (22%)[1] use a crossover design in which each subject is given a sequence of treatments to study within-patient differences or differences between cohorts until the crossover moment. Crossover trials, which can yield considerable cost savings, are statistically suitable for palliative or symptomatic treatment of chronic diseases and for single dose pharamocokinetic/pharmocodynamic studies. In cancer research however, a crossover design is used widely in a different way in studies that seek to establish basic efficacy of a novel agent. Cancer studies likely include the option to crossover to the experimental treatment to appeal to candidate patients and bolster trial recruitment, but also often out a sense of doing the right thing: maximizing the number of patients who have access to an investigational drug. Utilizing crossover designs without carefully considering the impact on scientific inference, however, may be harmful. Crossover designs can interfere with a study's ability to answer a clinical question, or worse, provide misleading estimates of a drug's true effect, possibly harming countless future patients whose treatment decisions are based on faulty or inadequate data.

Consider a recent crossover study of everolimus in advanced renal cell carcinoma (RECORD-1)[2]. Patients who had progressed on one or two prior lines of therapy were

Address correspondence to: Christine Grady, RN, PhD, Chief, Department of Bioethics, National Institutes of Health, Clinical Center, 10 Center Dr. Building 10/1C118, Bethesda MD 20892, CGrady@cc.nih.gov.

randomized, 2 to 1, to everolimus or placebo. Upon progression, placebo recipients were allowed to crossover to the study drug. The authors based their choice of a crossover design on, "both ethical and recruitment considerations.[2]" While the trial found a statistically significant difference in time to progression, there was no difference in overall survival, which the authors felt was, "probably due to confounding by crossover." The challenge here is that by consciously designing the trial with crossover, the investigators would have known that any uncertainty about a mortality benefit could be attributed to the confounding factor they introduced. Nevertheless, a post-hoc statistical analysis performed by the drug's manufacturer, used a statistical model to correct for crossover, and concluded that everolimus adds 0.93 days of life for every day it is taken[3].

The United Kingdom's (UK) National Institute for Health and Care Excellence (NICE) reviewed the data concerning everolimus, and ultimately did not recommend it[4]. The group felt that the estimate provided by the manufacturer's model was overly optimistic, and instead granted a smaller overall mortality benefit based on alternative statistical estimates. The group further cautioned, "any estimate of overall survival obtained using statistical modeling would be subject to some uncertainty because a number of assumptions would have to be made.[4]" NICE rejected everolimus for advanced renal cell carcinoma because, as the magnitude of its effectiveness was uncertain, so too were estimates of cost effectiveness, with several analyses finding cost-benefit ratios that exceeded NICE's limits for end of life treatments. But even NICE's analysis was forgiving, making a central favorable assumption regarding everolimus' efficacy (survival benefit was lost by crossover). There is at least one alternate interpretation of the RECORD-1 trial, i.e. that delaying treatment has no downside, and patients might benefit from waiting until progression to begin the drug. Patients in the placebo group were less likely to experience adverse events than those on everolimus, took the study drug for fewer days, and, did all this, while experiencing similar survival. So, why not wait?

Finally, let's consider yet another plausible interpretation from the current crossover investigation. Thus far, we have assumed a net beneficial survival effect from everolimus. But isn't it conceivable that everolimus, although it slows tumor growth and defers a handful of cancer deaths, modestly increases death for non-cancer reasons and has no net effect on overall mortality? Or that gains in PFS simply do not translate into improvements in OS? In either case, there would be a PFS benefit, but no change in overall survival. The drug would have real cost and side effects, but no real benefit. This is the ultimate worry regarding crossover study designs. The built-in confounding-by-design does not rule out that possibility.

Crossover designs, as frequently used in cancer studies, are different than the traditional form. A traditional crossover study occurs when both the treatment and control or placebo groups are, at some point in time, and typically after a "wash-out" period, switched to the other arm. Each patient can then serve has his/her own control, permitting powerful paired statistical testing to evaluate differences. As noted above, traditional crossover studies are used for single dose pharmacokinetic/pharmacodynamic studies, and for studies that utilize physiologic or subjective endpoints, which allow for multiple measurements over time. For instance, consider a recent crossover trial testing duloxetine among patients with chemotherapy induced peripheral neuropathy. This study fulfills all of the traditional requirements of crossover trials: it concerns a subjective endpoint, both groups were crossed over, and participants were compared to themselves[5]. In contrast, crossover is used in cancer trials typically to allow only the placebo or standard of care group to crossover and receive the experimental agent typically because their disease has progressed.[6] Additionally, cancer crossover trials typical examine tumor progression as a primary endpoint, but also consider overall mortality as a secondary endpoint. As such, some may consider crossover in cancer

trials to be a misnomer, as these studies may be more similar to a modified stepped wedge trial design[7]. To be clear, we confine our analysis only to crossover trials as they are frequently used in oncology research, and, in recent years, there has been no shortage of them.

Another recent trial[8] examined whether adding trametinib, a MEK inhibitor, to dabrafenib, a BRAF inhibitor, would benefit patients with metastatic melanoma. Patients receiving dabrafenib alone tended to progress a few months before those on dual therapy and were allowed to crossover upon progression. Overall mortality curves are indistinguishable. Again, three possibilities emerge: crossover masked an overall survival advantage; delaying the addition of trametinib confers no harms to patients; or, worst case scenario, trametinib slows dying from cancer but increases off target deaths. Small sample sizes and short follow up leave us uncertain of which explanation is right.

A careful and rigorous study design capable of answering a useful clinical or scientific question is an ethical requirement of clinical research. Participants of clinical research accept some risks justified by the importance of the knowledge to be gained. Ethically, benefits to participants should be maximized but should not compromise the scientific validity of a study. In an effort to offer the hope of benefit to ill participants in the control arm, oncology investigators often build in an option for participants who progress to cross over to the experimental arm. This carries the implicit assumption that the investigational agent is beneficial, subverting the concept of true clinical equipoise. However, empirical research[9] shows that less than half of novel agents tested in late stage oncology trials are beneficial, and far fewer (15%) are truly breakthroughs. The far more likely possibility is that by seeking to meet the hopes of participants and provide them with “benefit”, we undermine the ability of a study to answer the most important clinical question, overall survival. A study, which is designed in a way that it cannot answer a clinically meaningful question, is itself unethical, as it exposes all participants to risks without countervailing societal benefit. The crossover option in many studies confounds estimates of overall survival such that we do not know whether a treatment truly prolongs life.

Can this crossover strategy be improved or should it be abandoned? For clinical trials testing fundamental efficacy, unless and until a surrogate endpoint is shown to validly predict overall survival, it is hard to justify this choice of study design. A drug that slows tumor growth will improve progression free survival, but overall survival could be improved, unchanged, or worsened, and even a sophisticated statistical test cannot always untangle this hardwired bias. While short follow up and small sample sizes may complicate the interpretation of studies that use a crossover strategy, the crossover design itself is sufficient to obscure conclusions about survival. On the other hand, when efficacy has already been proven for a drug in a particular setting, crossover may be used meaningfully—as a test of treatment strategies: at what point does treatment benefit patients? This principle is at the heart of trials testing early versus delayed treatment for a given malignancy[10,11], and also similar to the question addressed in studies of maintenance chemotherapy[12]. While proponents of crossover may contend that trial recruitment will suffer without the promise of a crossover, we are aware of no studies that confirm crossover studies recruit faster. Additionally, 2:1 randomization is another technique claimed to improve recruitment, and one which is less likely to distort interpretations.

Crossover of patients from the control arm to the experimental arm after disease progression, as occurs in many cancer trials can be an obstacle to meaningful research. Most researchers employ it out of a sense of doing the right thing. And yet, in the life of a cancer drug there is limited chance to ensure that the benefits of treatment outweigh the risks. By utilizing crossovers in seminal trials used for applications for approvals of new drugs or new

indications, we preclude the ability of clinical trials to reach valid conclusions regarding survival benefits of therapies. These shortcomings might be tolerable if there was another opportunity to get it right. But, too often, there isn't. Doctors, patients, and payers see how long they can tolerate drugs like everolimus, managing real side effects, and hoping (but not knowing) if they are seeing real benefits.. "The challenge in oncology is to be sure that we remain focused on true clinical benefit—prolonging life."[13] Patients, particularly those at the end of life, deserve to know that the medications they use do more than affect imaging results; but that they actually improve the quality or quantity of life.

Acknowledgments

Acknowledgements and Disclaimer: We would like to thank Dr. Barry Kramer and Dennis Dixon for comments on earlier drafts of this paper and the anonymous reviewers for their valuable suggestions. The opinions expressed are the authors' own. They do not reflect any position or policy of the National Institutes of Health, U.S. Public Health Service, or Department of Health and Human Services.

References

1. Chan A-W, Altman DG. Epidemiology and reporting of randomised trials published in PubMed journals. The Lancet. 2005; 365:1159–62.
2. Motzer RJ, Escudier B, Oudard S, et al. Efficacy of everolimus in advanced renal cell carcinoma: a double-blind, randomised, placebo-controlled phase III trial. The Lancet. 372:449–56.
3. Korhonen P, Zuber E, Branson M, et al. Correcting overall survival for the impact of crossover via a rank-preserving structural failure time (RPSFT) model in the RECORD-1 trial of everolimus in metastatic renal-cell carcinoma. Journal of biopharmaceutical statistics. 2012; 22:1258–71. [PubMed: 23075021]
4. [Accessed June 6, 2013] Everolimus for the second-line treatment of advanced renal cell carcinoma. at http://publications.nice.org.uk/everolimus-for-the-second-line-treatment-of-advanced-renal-cell-carcinoma-ta219/consideration-of-the-evidence
5. Smith EM, Pang H, Cirrincione C, et al. Effect of duloxetine on pain, function, and quality of life among patients with chemotherapy-induced painful peripheral neuropathy: a randomized clinical trial. JAMA : the journal of the American Medical Association. 2013; 309:1359–67. [PubMed: 23549581]
6. Rimawi M, Hilsenbeck SG. Making Sense of Clinical Trial Data: Is Inverse Probability of Censoring Weighted Analysis the Answer to Crossover Bias? Journal of Clinical Oncology. 2012
7. Brown CA, Lilford RJ. The stepped wedge trial design: a systematic review. BMC Med Res Methodol. 2006; 6:54. [PubMed: 17092344]
8. Flaherty KT, Infante JR, Daud A, et al. Combined BRAF and MEK Inhibition in Melanoma with BRAF V600 Mutations. New England Journal of Medicine. 2012; 367:1694–703. [PubMed: 23020132]
9. Djulbegovic BKASHP, et al. Treatment success in cancer: New cancer treatment successes identified in phase 3 randomized controlled trials conducted by the national cancer institute–sponsored cooperative oncology groups, 1955 to 2006. Archives of internal medicine. 2008; 168:632–42. [PubMed: 18362256]
10. Dighiero G, Maloum K, Desablens B, et al. Chlorambucil in Indolent Chronic Lymphocytic Leukemia. New England Journal of Medicine. 1998; 338:1506–14. [PubMed: 9593789]
11. Mateos M-V, Hernández M-T, Giraldo P, et al. Lenalidomide plus Dexamethasone for High-Risk Smoldering Multiple Myeloma. New England Journal of Medicine. 2013; 369:438–47. [PubMed: 23902483]
12. Kluin-Nelemans HC, Hoster E, Hermine O, et al. Treatment of Older Patients with Mantle-Cell Lymphoma. New England Journal of Medicine. 2012; 367:520–31. [PubMed: 22873532]
13. Amiri-Kordestani L, Fojo T. Why Do Phase III Clinical Trials in Oncology Fail so Often? Journal of the National Cancer Institute. 2012; 104:568–9. [PubMed: 22491346]