Perspective

Clinical Trial Design in Contemporary Device Studies in Heart Failure: Is There a Gold Standard?

PAUL J. HAUPTMAN, MD,[1] AND STEPHEN S. GOTTLIEB, MD[2]

St Louis, Missouri; and Baltimore, Maryland

ABSTRACT

The assessment of the efficacy and safety of implantable cardiac devices used for the management of heart failure is complicated by procedural challenges. We present an overview of the advantages and disadvantages of different clinical trial designs, and discuss investigator and patient blinding. We conclude that blinding is optimal, but methodologically difficult. Until rules for and assessment of blinding are developed or surrogate measures are considered to be acceptable from a regulatory standpoint, an open-label design with objective end points is an unavoidable default standard. (*J Cardiac Fail 2014;20:223–228*)
Key Words: Clinical trials, implantable devices, heart failure.

Major advances in heart failure therapeutics over the past 2 decades have involved both pharmacologic and device interventions. Randomized double-blind placebo-controlled trials with combined hospitalization and mortality end points have been the standard for phase III drug studies and are credited with proving efficacy and convincing physicians of a medication's worth. In contrast, proving efficacy in device trials is more problematic. "Mega-studies" of devices are generally cost prohibitive, and traditional blinding (or, per the United States Food and Drug Administration [FDA], masking) may not be feasible. With the known large effects of placebo, this can make determination of a device's value difficult. In this paper, we review trial designs and discuss contemporary challenges encountered when designing clinical device studies and interpreting the results from both a scientific and regulatory framework. In particular, we try to answer the question, "Is there a gold standard design for device studies?," a crucial topic for the future of device innovation and clinical development.

The Effects of Placebo: Experience From Drug Trials

Numerous studies show the powerful effect of placebo in randomized drug studies. A retrospective evaluation of patients with heart failure who received placebo compared with patients who were part of a "natural control" showed a statistically significant improvement in exercise time and functional class in those who received placebo.[1] Because patient expectations appear to be very important, the quality of the placebo is also influential. For example, intravenous placebo lowered blood pressure compared with oral placebo in a randomized study.[2] Even the color of a pill has effects on efficacy.[3] It is therefore to be expected that a surgical procedure could have a very large placebo effect. To account for this, assessment of efficacy and safety of a device intervention would ideally entail the implantation of the device in all study subjects and double-blinded activation. However, as we discuss, this solution is often not feasible or practical.

Standard Approaches: Does the Wheel Need to Be Reinvented?

Randomized 2-armed studies involve either (1) a parallel open-label design in which patients are randomized to

From the [1]Department of Medicine, Saint Louis University School of Medicine, St Louis, Missouri and [2]Division of Cardiology, University of Maryland School of Medicine, Baltimore, Maryland.
Manuscript received November 4, 2013; revised manuscript received December 23, 2013; revised manuscript accepted January 10, 2014.
Reprint requests: Paul J. Hauptman, MD, Division of Cardiology, Saint Louis University Hospital, 3635 Vista Avenue, St Louis. MO 63110. Tel: 314 268 5293; Fax: 314 577 8861. E-mail: hauptmpj@slu.edu
See page 227 for disclosure information.
1071-9164/$ - see front matter

http://dx.doi.org/10.1016/j.cardfail.2014.01.009

"standard of care" (or optimal medical therapy) plus implant with activation versus a "standard of care" plus no implant comparator arm, or (2) a parallel-implant design in which all patients undergo implantation but activation occurs in one arm only (Fig. 1). In both designs, study subjects are monitored according to their respective treatment assignments until study completion defined by a prespecified follow-up period or number of events. In most cases, the parallel open-label trial requires hard end points, including mortality. Both patient and investigator know the treatment assignment and this knowledge influences clinical decision making and the thresholds for hospitalization. Patients may also have a natural bias in favor of (experimental) device therapy and out-migrate from the study when randomized to a standard of care arm. This phenomenon can occur despite the concerted efforts of the research team to maintain randomization and treatment assignment. In a parallel-implant design, the study can be either single blind (patient only) or double blind (patient and investigator). However, the patient who receives an implant without activation is subjected to the medical risks of the implant procedure without any chance of near-term benefit.

Do Crossover Designs Work?

Double Crossover

To account for medical and ethical concerns related to a trial that subjects all study participants to the risks of the implant procedure and to further assess device efficacy, a double-crossover design has been used. A typical example is the MUSTIC (Multisite Stimulation in Cardiomyopathies) study of cardiac resynchronization therapy, in which all patients were implanted but one-half proceeded with the device in an inactive state for 12 weeks before crossing over to activation; the others were deactivated after a similar duration of active therapy.[4]

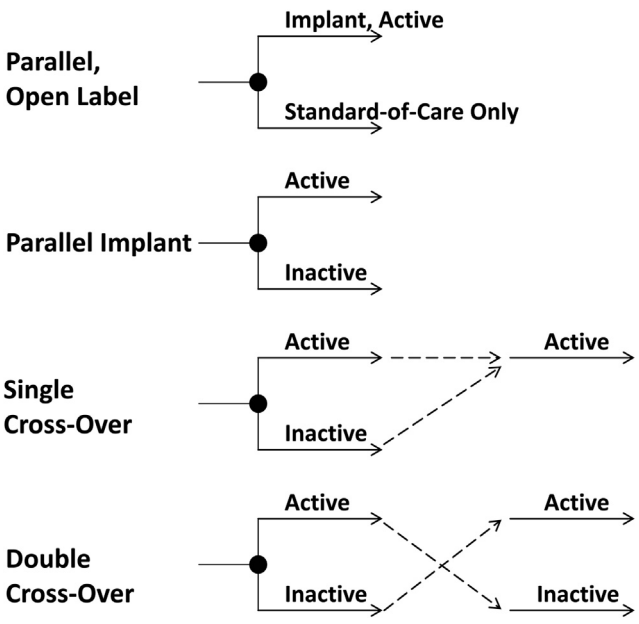

Fig. 1. Schema of study designs (see text). All study arms receive standard of care therapy (optimal medical management) regardless of treatment assignment.

A theoretical statistical concern for this design relates to the possibility that a therapeutic effect persists beyond the time of device deactivation. Such a phenomenon would lessen the likelihood of detecting a difference in outcomes between the previously active arm of the study and the newly activated arm that had been initially randomized to an inactive status, thereby affected statistical power.

Single Crossover

Few contemporary studies have repeated the MUSTIC double-crossover approach. One alternative, as described in the Rheos Pivotal Trial of baroreceptor stimulation,[5] is to perform device implantation in all subjects and use a modified crossover design in which those study subjects who were initially randomized to active therapy stay on active therapy while those who were initially not activated cross over (Fig. 1). This design allows for an extended time frame for observation of both safety and efficacy of active therapy while also permitting a comparison between active and inactive modalities. From a practical level, patients may be more willing to participate in a study if they know that once activated, they will continue to get the experimental therapy.

Study Duration in Crossover Designs: How Long Is Long Enough?

Whether the modified crossover design is optimal is unclear. A single crossover provides only 1 time-limited period for comparison between active and inactive device status. Ideally, if the timing of maximal therapeutic benefit were known before initiation of the pivotal study, the timing of the crossover could be carefully estimated. However, this is not typically known in advance. Theoretically, it would be possible to further assess efficacy in the crossover arm in a pre-post comparison, though this additional analysis is fraught with significant practical and statistical limitations given the progressive nature of heart failure (eg, event rates may accelerate over time).

In most studies, the crossover occurs in a relatively short time frame, which translates into a limited duration of follow-up and potentially inadequate time for a meaningful clinical therapeutic effect of active intervention to take place. For example, in an early comparison of cardiac resynchronization therapy (CRT) with implantable cardioverter defibrillator (ICD) versus ICD alone (CRT not activated), the original investigational plan included a 3-month crossover but was modified during the study owing to "regulatory concerns over … the length of follow-up."[6] Although the precise reasons for removal of the crossover were not delineated, it may be that 3 months were deemed to be insufficient to establish and confirm durability of clinical effect of CRT.