Available online at www.sciencedirect.com

ScienceDirect

journal homepage: www.JournalofSurgicalResearch.com

Association for Academic Surgery

Maturation of effect size during enrollment of prospective randomized trials

*Ashwini S. Poola, MD, Tolulope A. Oyetunji, MD, George W. Holcomb III, MD, and Shawn D. St. Peter, MD**

Department of Pediatric Surgery, Children's Mercy Hospital, Kansas City, Missouri

ARTICLE INFO

Article history:
Received 9 February 2017
Received in revised form
17 May 2017
Accepted 19 June 2017
Available online xxx

Keywords:
Effect size
Sample size
Randomized control trial

ABSTRACT

Background: Randomized clinical trials are powered by calculating the minimum sample size required to achieve statistical significance, given an estimated effect size (ES). The ES is the raw difference between two treatment arms. ES quantifies the actual magnitude of clinical differences between cohorts and is usually reflective of the true meaning of the trial, regardless of statistical significance. Under a fixed protocol, we hypothesize that the ES may be attained at a smaller sample than predesigned. To investigate patterns of ES during enrollment, we analyzed completed trials that were completed at our institution.

Methods: Outcomes of 11 prospective randomized clinical trials from our institution were reviewed. ES was calculated at intervals throughout each trial to determine at which point a steady clinical difference was achieved between treatment cohorts.

Results: ES stabilized at a median of 64% enrollment. All patients were needed to meet the precise ES in our smallest study, indicating the need for full enrollment in smaller studies. Otherwise, 50% of our trials required between 48% and 76% of patient enrollment to meet ES. In comparing clinical outcomes, 9 of 12 found a final difference that was nearly identical to the difference that could have been determined much earlier. Categorical outcomes met stabilized ES at 51% enrollment and continuous outcomes at 68%.

Conclusions: ES and final clinical outcomes were achieved before the completion of enrollment for most of our studies. This suggests that clinical differences detected by randomization may not necessarily require the robust sample size often needed to establish statistical significance. This is particularly relevant in fixed-protocol interventional trials of homogenous populations.

Introduction

Sample size calculation is the foundation for randomized controlled trials. It is generally accomplished by starting with an estimate of the difference anticipated to be observed between two treatment options under the trial protocol.[1,2] Given other parameters such as type I error and desired power, the sample size is the number of patients needed to detect a statistically significant difference in the variable established as the primary outcome. In all types of comparative clinical studies, achieving statistical significance has become the focus of concern, perhaps to the point of devaluing not just the

* *Corresponding author*. Department of Pediatric Surgery, Children's Mercy Hospital, 2401 Gillham Road, Kansas City, MO 64108. Tel.: +1 816 983-6465; fax: +1 816 983-6885.
E-mail address: sspeter@cmh.edu (S.D. St Peter).

0022-4804/$ – see front matter © 2017 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.jss.2017.06.082

presence of differences but also the reproducibility of these differences in treatment cohorts.[3] Statistically powering a study to determine a significant difference is not equivalent to determining a meaningful effect. Perhaps, more important than identifying the presence or absence of significance is to identify the magnitude of difference between two arms given a trial protocol. This magnitude is the effect size (ES).[4]

Estimation of ES is one of the key steps in a priori calculations to appropriately power a study to meet significance.[3] However, understanding ES may prove useful in other aspects of study design and monitoring. There are no reports in the surgical literature evaluating ES maturation during enrollment of randomized trials. Therefore, we reviewed the raw differences of primary outcomes from our completed trials throughout enrollment to discern when ES stabilized. We hypothesize that, under a fixed protocol with a homogenous population, stable ES may be achieved well before the number of patients intended to identify a statistically significant difference.

Materials and methods

Primary outcomes of completed prospective randomized controlled trials were reviewed at our institution. All trials were conducted with strict management protocols. Sample size for each of these trials was calculated by using a significance level of 0.05 and a power of 0.8 or 0.9. The findings of these studies have all been previously published.[5-15] Standard comparative tests (P value) were used to determine significance or lack thereof, but ES was neither calculated nor was described.

The ES was calculated as the raw mean difference between two cohorts of each trial at short intervals throughout each trial. This was trended to determine the point in each study that a steady difference was achieved between treatment cohorts. Raw outcomes at this stabilization point were compared with outcomes at the completion point of the study to evaluate those differences that may occur over the course of a trial recruitment period.

Results

Twelve outcomes were reviewed from eleven completed trials (we included one secondary outcome in the analysis). Trials 3, 6, 9, and 10 studied surgical techniques of standard pediatric operations: cholecystectomy, appendectomy, fundoplication, and pyloromyotomy. The remaining seven studies examined management strategies in children with empyema, perforated appendicitis, pectus excavatum, and burns. Table 1 lists the specific outcomes that were reported for each trial.

Seven of the studies evaluated continuous outcomes: trials 1, 3, 4, 5, 6A, 7, and 10. The remaining five outcomes were categorical. Of these, eight were found to have a final difference that was near identical to a raw difference that could have been determined earlier, exclusive of trial 1, which required full recruitment of study patients to achieve ES. Categorical outcomes stabilized at 51% enrollment and continuous outcomes at 68%.

The completed trials, with the calculated sample size and the number of enrolled patients at which ES stabilized, are listed in Table 2. The ES plateaued at a median enrollment of 64% of original sample size of patients. All of the patients were needed to meet the precise ES of 0.1 in our smallest study, trial

Table 1 – Description of clinical trials.

Trial	Cohort A	Cohort B	Study population	Outcome
Trial 2[5]	Video-assisted thoracoscopic surgery	Fibrinolysis	Empyema management	Hospitalization (days)
Trial 2[6]	Suction	Irrigation	Perforated appendicitis intra-operative management	Abscess rate (%)
Trial 3[7]	Four-port laparoscopic cholecystectomy	Single-incision laparoscopic cholecystectomy	Cholecystectomy technique	Operative time (minutes)
Trial 4[8]	Tissue plasminogen activator instillation	Normal saline instillation	Perforated appendicitis post-operative abscess drain management	Hospitalization (days)
Trial 5[9]	Double antibiotics	Triple antibiotics	Antibiotic regimen in perforated appendicitis	Hospitalization (days)
Trial 6A[10]	Three-port appendectomy	Single-incision laparoscopic appendectomy	Appendectomy technique	Operative time (minutes)
Trial 6B[10]	Three-port appendectomy	Single-incision laparoscopic appendectomy	Appendectomy technique	Wound infection rate (%)
Trial 7[11]	Epidural	Patient-controlled analgesia	Pectus excavatum post-operative Pain management	Hospitalization (days)
Trial 8[12]	Silver sulfadiazine	Collagenase	Burn: wound management	Graft rate (%)
Trial 9[13]	Maximal dissection	Minimal dissection	Fundoplication technique	Hernia rate (%)
Trial 10[14]	Open	Laparoscopic	Pyloromyotomy technique	Operative time (minutes)
Trial 11[15]	Intravenous for 5 d	Home oral antibiotics	Antibiotic regimen in perforated appendicitis	Abscess rate (%)