See the corresponding article in this issue, pp 111–117.

J Neurosurg Pediatrics 2:109–110, 2008

Editorial

It's randomized and double blinded . . . what more do we want?

JOHN R. W. KESTLE, M.D.

Division of Pediatric Neurosurgery, Primary Children's Medical Center, Salt Lake City, Utah

This issue of *Journal of Neurosurgery: Pediatrics* presents a randomized trial in which investigators have evaluated antimicrobial suture (AMS) in the prevention of shunt infection. The authors randomized 84 shunt procedures in 61 patients over 21 months. The surgeons and patients were blinded to treatment group. The groups appeared to be balanced with respect to shunt infection risk factors. Infection within 6 months of surgery occurred in 2 (4.3%) of 46 AMS procedures and in 8 (21%) of 38 control procedures.

This study is a good first step. The authors should be applauded for conducting a double-blinded randomized trial, but the results need to be considered preliminary, and as they stand are not sufficient for a change in practice. As the authors state, they need further evaluation in a larger randomized trial.

Why? What more do we need before we adopt AMSs for shunt surgery? There are a number of issues that should be addressed in a definitive trial. The authors appropriately recognize some of these in their manuscript. I would like to highlight them.

1) The infection rate in the control group in the study was high (21%). This is a function of a small sample size. With only 38 patients in the control group, 1 or 2 events will have a large impact on the results. If the infection rate in the control group had been closer to their usual (9%), a much larger sample size (1062 patients) would have been required. This issue can certainly be addressed in a larger trial.

2) A study hypothesis should be defined ahead of time. It should specify the difference that the investigators are interested in detecting (that is, cutting their infection rate in half or by x%) and have adequate power to detect the smallest difference that is clinically important. This may have been done, but it was not described in the report. As a result it appears that the authors kept randomizing patients until they got a "significant" difference.

3) Patients were "rerandomized." In this case the number was small and it is unlikely to have a large impact on the results, but I would advise against this in the definitive trial. The analysis assumes that each patient entered is independent, and this is not true when patients are randomized more than once.

4) The technique of cerebrospinal fluid (CSF) culture was not described. Were cases always kept long enough to identify *Propionibacterium acnes*?

5) Because the outcome (positive culture) is dependent on the results of a shunt tap, the indications for doing a tap should be clearly defined.

6) They report an interim analysis. At the first interim analysis the results were "not significant," so they continued accrual. At the time of a second interim analysis there was a significantly higher infection rate in the control group, and new patient enrollment was halted. I find this puzzling. An interim analysis in a clinical trial affects the sample size, and there are well-defined techniques for doing this.[1] Usually the results have to be extreme to justify early stopping—so extreme that the investigators (or more commonly an independent data safety monitoring board) feel uncomfortable randomizing any more patients to the control arm of the study. Apparently the authors felt this way and stopped the study, and yet they are now proposing another larger trial. In light of these results and their decision to stop, can they randomize patients in the future and convince others to do so?

7) "Patient population characteristics did not differ significantly. . . ." As so many authors do, this paper includes probability values beside each of the baseline parameters to test randomization. These values give the probability that the observation occurred by chance. In a randomized trial, the observation (distribution of baseline parameters between the groups) was due to chance, by definition. All probability values are 1.0. In addition, there is no power calculation for such a test, so a "nonsignificant" probability value is meaningless. The table is appropriate (the probability values are not). Factors that appear to be imbalanced should be evaluated based on their clinical importance, and if an important imbalance exists it should be addressed in the analysis.

Overall I am delighted to see this paper: we need more like this! I congratulate the authors on a well-done pilot study, but we cannot use this paper to justify the use of AMSs in shunt surgery. I agree with the authors that we must use their work to plan a definitive trial.

Reference

1. Friedman LM, Furberg CD, DeMets DL: Monitoring response variables, in **Fundamentals of Clinical Trials, ed 2.** Littleton, MA: PSG Publishing Co., 1985, pp 213–240

Response

Curtis J. Rozzelle, M.D.
Jody Leonardo, M.D.
Veetai Li, M.D.

Women and Children's Hospital of Buffalo, Kaleida Health, School of Medicine and Biomedical Sciences, University at Buffalo, State University of New York, Buffalo, New York

We appreciate Dr. Kestle's thoughtful analysis, kind remarks, and constructive critique of our small single-center trial. We do not claim to present a definitive study, merely to report interesting and potentially beneficial findings related to the problem of shunt infection. It is our hope that these findings will generate sufficient interest within the pediatric neurosurgical community for the definitive study to be designed, implemented, and completed.

As stated, some issues were elucidated in our manuscript. The common theme for these is that compromises were made to accrue procedures quickly from a relatively modest-sized pediatric neurosurgical service. We offer the following comments to amplify and further explain these issues.

1) The control group infection rate was higher than our institutional "baseline." Continuing the study at a single center, however, would have taken a prohibitive amount of time to accrue > 1000 procedures.

2) With this limitation in mind, the study was initiated without power calculations. Every conservative power estimate generated beforehand indicated that the trial could not be completed in < 5 years. The study design was kept simple and inexpensive to minimize the "investment."

3) The decision to "rerandomize" patients was made to enroll as many shunt procedures as possible, accepting that this design aspect is less than ideal.

4) The microbiology laboratory at our institution has standing orders to monitor all CSF cultures from shunts for *P. acnes*, and all CSF shunt aspirates are sent for aerobic and anaerobic cultures.

5) At our institution all patients with shunts who present with symptoms/signs of shunt malfunction and/or shunt infection within 6 months of any operative shunt procedure undergo shunt tapping to rule out infection.

6) After 40 procedures had been randomized and the patients followed for 6 months, the control group had 4 infections and the AMS group had none. This difference yielded a right-tailed probability value of exactly 0.05; merely a trend. Randomization continued until the results reported were observed, although not all patients had reached 6 months postprocedure. At that point enrollment was halted and no additional infections were detected. We were uncomfortable randomizing more patients primarily because the control group's infection rate was so high. We believe a larger trial is necessary and appropriate due to the study limitations previously enumerated.

7) We stand corrected. Lesson learned.

In conclusion, we would be delighted to collaborate with Dr. Kestle and others who would be willing to participate in a definitive trial. *(DOI: 10.3171/PED/2008/2/8/109)*