J Neurosurg (Pediatrics 2) 101:130–133, 2004

Assessing the validity of the endoscopic shunt insertion trial: did surgical experience affect the results?

Paul Klimo Jr., M.D., M.P.H., Clinton J. Thompson, B.S., James Drake, M.B., B.Ch., and John R. W. Kestle, M.D., M.Sc.

Departments of Neurosurgery and Family and Preventive Medicine, University of Utah, Salt Lake City; Division of Pediatric Neurosurgery, Primary Children's Medical Center, Salt Lake City, Utah; and Division of Neurosurgery, University of Toronto, The Hospital for Sick Children, Toronto, Ontario, Canada

Object. Most surgical procedures are associated with a learning curve in which the success rate is lower early in the experience before mistakes have been identified and modifications made to the procedure. Negative results obtained early in a trial's learning curve may be a matter of timing rather than a reflection of the procedure's effectiveness. The recently published results of the Endoscopic Shunt Insertion Trial (ESIT) represent the notion that endoscopically placed shunts were no more likely to survive than conventionally placed shunts. This negative result may be due to inexperience in performing endoscopic surgeries.

Methods. Surgical experience was assessed in two ways. Shunt survival rates were compared between cases treated endoscopically in the 1st and last years of the ESIT. The effect of center volume was evaluated using a Cox proportional hazard model in which the following variables were analyzed: age at registration, the diagnosis of myelomeningocele, head size, method of shunt placement (endoscopic compared with conventional), and center volume.

There was no difference in survival (endurance) of the shunt between patients enrolled in the 1st and last years (log rank = 0.08, p = 0.77). Likewise, no variable in the Cox multivariate model, including center volume, was a significant factor in predicting shunt survival.

Conclusions. The primary result of the ESIT was found to be internally valid. The fact that endoscopic shunt placement did not benefit patients evaluated in the study was not due to early timing of the trial. Any learning curve among the participating surgeons did not adversely affect the results.

Key Words • endoscope • ventriculoperitoneal shunt • validity • randomized clinical trial • pediatric neurosurgery

Hydrocephalus is arguably the most common disorder in pediatric neurosurgery. Treatment usually involves placement of a shunt system, most commonly a ventriculoperitoneal one. Recently, an RCT was completed in which endoscopic-assisted placement of the ventricular catheter was compared with the conventional placement method.[15] The main outcome was shunt failure, defined as obstruction, overdrainage, loculation, or infection. The investigators found no statistically significant benefit to using an endoscope to place the ventricular catheter. Because many new surgical procedures, including endoscopic shunt placement, are associated with a learning curve, we considered the possibility that the lack of benefit in patients who underwent endoscopic shunt surgery may have been the result of starting the trial too early, while surgeons were still learning the procedure.

The RCT is generally accepted as the best trial design for determining the efficacy of therapeutic (medical or surgical) interventions. Because of its rigorous design and concurrent control group, it minimizes error (both random and systematic, otherwise known as bias) and tends to distribute all known and unknown prognostic factors (confounders) equally between the treatment groups. Thus, any difference in outcome between the treatment groups is more likely a result of the intervention being investigated. When designing and evaluating the results of an RCT in which a surgical procedure is evaluated, each surgeon's experience with the technique must be considered. For example, in the EC–IC Bypass Study,[9,10] surgeons were required to demonstrate the achievement of at least 80% graft patency in at least 10 consecutive EC–IC procedures. In this way, the designers of the trial controlled for surgical inexperience as a potential confounder prior to the initiation of the study. If it is not possible to control for surgical experience prior to beginning the trial, this variable should be analyzed during and at the conclusion of the trial. Often both methods are used to assess surgical experience.

The timing of trials is an important issue. All surgeons would agree that innovative, new surgical procedures are associated with a learning curve, a phenomenon that reflects the change in success rate over time. Mistakes and

Abbreviations used in this paper: EC–IC = extracranial–intracranial; ESIT = Endoscopic Shunt Insertion Trial; RCT = randomized clinical trial.

modifications involving the procedure may occur during the early part of the learning curve until the procedure has been sufficiently refined. The slope and length of different portions of the learning curve primarily depend on a surgeon's experience and the complexity of the procedure itself as well as other factors such as improved technology and facilities. If a clinical trial is started at the beginning of the curve, a false-negative result may be obtained simply because the surgeons have not yet perfected the procedure. In the ESIT, participating surgeons were required to have performed five or more endoscopic shunt procedures; however, this may not have assured technical proficiency. Therefore, it is possible that the shunt survival rate at the beginning of the trial was low simply because surgeons had to learn how to use the endoscope. To investigate this possibility, we hypothesized that if the trial was initiated before the participating surgeons had acquired adequate experience with the endoscope, we might observe evidence of "learning" during the trial. In other words, if the surgeons were learning during the trial, we would find better results at higher-volume centers and/or improved results as the course of the accrual period progressed. Absence of either of these phenomena would indicate that the trial was not conducted during the learning phase for the procedure and that the negative result is valid.

Clinical Material and Methods

The detailed methods of the ESIT have been published previously.[15] The accrual period for the trial was 3.5 years. To test our hypothesis, demographic data were collected for patients randomized to the endoscope-treated group during the 1st and last years of the trial. Data included age at registration (in days), sex, whether the patient was born with a myelomeningocele, head circumference at time of registration (in centimeters), and the mean shunt survival (in years). Kaplan–Meier survival curves pertaining to the shunts themselves were calculated for each treatment group. To determine whether there was a learning curve for the endoscope, the trial accrual period was divided into seven 6-month intervals. One-year shunt survival rates for each of these intervals were calculated.

The effect of the center's treatment volume was evaluated by constructing a Cox multivariate proportional hazard regression model. Variables assessed in the model were age at registration, the presence of a myelomeningocele, head size at time of registration, whether the shunt was placed with the aid of an endoscope, and the center's treatment volume (number of study participants provided by that center). Age, head circumference, and center volume were kept as continuous or discrete variables. Statistical significance occurred at probability levels less than 0.05. All statistical calculations were conducted using Stata software (version 8; Stata Corp., College Station, TX).

Results

The results of the ESIT will be briefly summarized. Eligible patients were younger than 18 years of age, exhibited clinical and radiographic evidence of hydrocephalus, and required a cerebrospinal fluid diversion procedure. During a 3.5-year period, 393 (214 male and 179 female) patients were randomized: 194 to the endoscopic treatment arm and 199 to the conventional insertion arm. The primary end point was shunt failure, defined as obstruction, overdrainage, loculation, or infection. The overall 1- and 2-year incidences of shunt failure were 0.38 and 0.47, respectively. The 1-year shunt survival was 0.58 in the endoscope group and 0.66 in the nonendoscope group (log rank = 2.92, p = 0.09).[17]

TABLE 1

Demographic data obtained in patients randomized to the endoscope treatment arm for the 1st and last years of the ESIT

Variable	Period in the Study (%)	
	1st Year	Last Year
no. of cases	32	61
male	17 (53.3)	27 (44.3)
median age at registration (days)	71.5	72
myelomeningocele	13 (40.6)	15 (24.6)
median head circumference (cm)	40.9	40.5
median shunt survival (yrs)	1.22	0.93

The demographic data obtained in patients randomized in the 1st and last years of the trial are summarized in Table 1. There were 32 patients in the 1st year compared with 61 in the last year. The two cohorts were similar in terms of age, sex, and head circumference at registration, although there were more patients harboring a myelomeningocele in the 1st-year cohort. The median 1-year shunt survival rate for the 1st- and last-year patients were 1.22 and 0.93, respectively.

There was no significant difference in shunt survival in cases in which the endoscope was used when 1st- and last-year data were compared (log rank = 0.08, p = 0.77; Fig. 1). The 1-year shunt survival rates for each of the 6-month intervals are shown in Fig. 2. There was no sustained improvement in the survival rates during the study. The results of the Cox proportional hazard model analysis are shown in Table 2. No factor, including center volume, was found to be statistically significant. Only head circumference was indicative of predicting shunt failure (p = 0.053).

Discussion

Clinical trials in which a new surgical procedure is being evaluated are more complex to design than the typical medical or drug trial for various reasons.[13,17,26,30,31] Standardizing the procedure of interest may be difficult, especially as the complexity of the procedure increases. Blinding is often impossible. There are ethical issues unique to surgical trials, such as allowing chance to determine whether a patient will undergo an invasive procedure and offering a placebo or sham surgery. Accruing patients in surgical trials is more challenging. Experience and learning associated with the surgical intervention and their relationship to the timing of the trial must be taken into account. These and other components of an RCT are important enough that investigators often present and publish their methodology prior to initiating the trial.[10,18,22]

Most procedures have an associated learning curve, which can have significant impact on patient care.[12,29] In general, a successful procedure-related outcome is more likely after mistakes have been corrected and modifica-

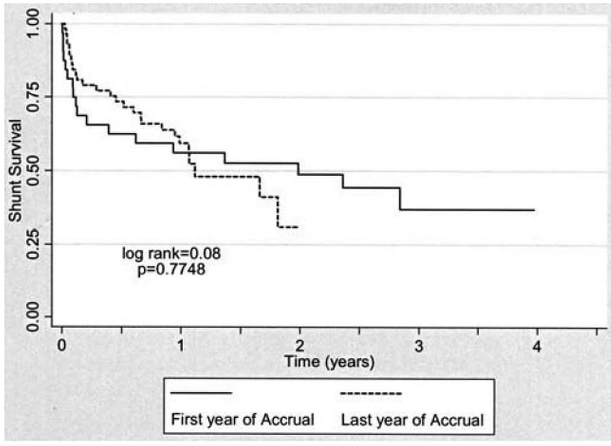

FIG. 1. Graph showing Kaplan–Meier curves. The shunt failure rate in 93 endoscope-treated patients accrued during the 1st year *(solid line)* and last year *(dotted line)* of the trial are depicted.

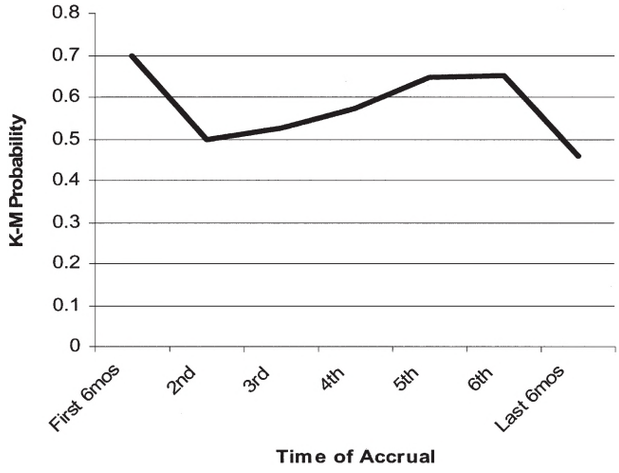

FIG. 2. Graph demonstrating 1-year shunt survival rates for each of the seven 6-month intervals of the trial. K-M = Kaplan–Meier.

tions made. Learning curves have been investigated for many new procedures such as endovascular repair of abdominal aortic aneurysms,[16] breast cancer lymphatic mapping,[1] laparoscopic colectomy[28] and cholecystectomy,[20] transurethral ultrasonography-guided laser-induced prostatectomy,[27] and sutureless intestinal anastomosis.[8] There are numerous examples within the neurosurgical literature that underscore a positive volume and experience–outcome relationship.[2,5,25] For example, a carotid endarterectomy confers a protective benefit in patients with stenosis greater than 60% only if the perioperative combined morbidity and mortality rates are less than 3 and 6% for asymptomatic and symptomatic patients, respectively.[11,19,21] Mortality rates were found to be lower in higher-volume centers when patients underwent surgery for ruptured and unruptured aneurysms.[6,7] Ramsay, et al.,[23] published a comprehensive review in which they described how learning curves have been assessed for emerging health technology. They found 272 studies that met their inclusion criteria. In 51% of the studies minimally invasive surgeries were assessed, and 95% of the articles were case series. In only 2% of the studies were data collected from RCTs. The authors found that the statistical methods used to assess learning have almost always been crude, such as descriptive data without formal statistical testing. In fact, Ramsay, et al.,[24] found that learning curves described in nonmedical fields such as engineering and psychology often provided more sophisticated methods.

The presence or absence of a learning curve and its effect on the timing of the trial is important. Chalmers strongly believed that " . . . from the scientific, ethical, and practical standpoint, exploration of any new therapy in sick patients should begin with randomization into either the conventional or the new treatment regimen."[4] If the trial is conducted too early, however, success rates may be artificially suppressed simply because participating surgeons are unfamiliar with the procedure and are still "learning." Some authors have argued that surgical trials should be conducted only after the pertinent intervention has been perfected (or as nearly as possible) and the slope of the learning curve is nearing zero. By the time the procedure has been perfected, however, clinical equipoise may have disappeared as people become convinced of the procedure's effectiveness based on poor evidence. Such a paradox is known as the Buxton law, in which the following is stated: "It is always too early for rigorous evaluation until, unfortunately, it's suddenly too late."[3]

Based on our analyses, we found no evidence to indicate that the participating surgeons were learning during the trial. The survival of the shunt was similar for endoscope-treated patients accrued in the 1st and last years of the trial. This finding was true despite the fact that myelomeningocele is a risk factor for shunt failure and there were more patients with myelomeningocele entered in the 1st year of the study. Furthermore, center volume was not found to be predictive of shunt survival in the Cox multivariate proportional hazard model. Interestingly, head circumference was the only variable that may be related to shunt failure. Thus, we believe that the results of the ESIT are valid; the negative result cannot be attributed to lack of surgeon experience or the trial being performed too early.

When designing a trial in which surgical experience and learning could potentially distort the true effect of the surgical intervention, the investigators should be aware of the procedure's learning curve and its potential impact on the outcome of the trial. Based on published and pilot data (such as a prerandomization phase of observational data collection), it may be possible to determine whether the

TABLE 2

Summary of data derived from the Cox multivariate proportional hazard model

Variable	Hazard Ratio	p Value
age at registration	1.00	0.57
myelomeningocele	0.83	0.74
shunt placement	1.19	0.29
center volume	1.00	0.74
head circumference	0.97	0.05

results of a particular procedure are stable or improving over time.[23] The trial design may minimize the potential for learning by imposing rigid entry criteria for surgeons as was done in the EC–IC bypass trial.[9] Alternatively, the presence and impact of learning can be assessed in a post hoc fashion as in the Shunt Design Trial.[14] In that trial, various shunt systems were evaluated; the primary outcome was shunt failure. Because two of the valves were new at the time of the study, it was possible that the surgeons had to learn how to place them and this learning process might have affected the results. Shunt failure was subsequently analyzed both chronologically (early compared with late phases of the trial) and in terms of center volume. As in our analysis, learning did not affect the results. nevertheless, failure to account for a learning curve may lead to inaccurate study results in some instances and, potentially, the abandonment of a beneficial procedure.

Conclusions

The failure of endoscopic assistance to improve shunt survival in the ESIT was not due to inexperience with the endoscope. No appreciable learning curve existed, and center volume had no effect on outcome.

References

1. Bass SS, Cox CE, Reintgen DS: Learning curves and certification for breast cancer lymphatic mapping. **Surg Oncol Clin N Am 8:**497–509, 1999
2. Buchman CA, Chen DA, Flannagan P, et al: The learning curve for acoustic tumor surgery. **Laryngoscope 106:**1406–1411, 1996
3. Buxton MJ: Problems in the economic appraisal of new health technology: The evaluation of heart transplants in the UK, in Drummond MF (ed): **Economic Appraisal of Health Technology in the European Community.** Oxford: Oxford University Press, 1987, pp 103–118
4. Chalmers TC: Randomization of the first patient. **Med Clin North Am 59:**1035–1038, 1975
5. Cochrane DD, Kestle JRW: The influence of surgical operative experience on the duration of first ventriculoperitoneal shunt function and infection. **Pediatr Neurosurg 38:**295–301, 2003
6. Cowan JA Jr, Dimick JB, Wainess RM, et al: Outcomes after cerebral aneurysm clip occlusion in the United States: the need for evidence-based hospital referral. **J Neurosurg 99:**947–952, 2003
7. Cross DT III, Tirschwell DL, Clark MA, et al: Mortality rates after subarachnoid hemorrhage: variations according to hospital case volume in 18 states. **J Neurosurg 99:**810–817, 2003
8. Dyess DL, Curreri PW, Ferrara JJ: A new technique for sutureless intestinal anastomosis. A prospective, randomized, clinical trial. **Am Surg 56:**71–75, 1990
9. The EC/IC Bypass Study Group: Failure of extracranial-intracranial arterial bypass to reduce the risk of ischemic stroke. Results of an international randomized trial. **N Engl J Med 313:** 1191–1200, 1985
10. The EC/IC Bypass Study Group: The International Cooperative Study of Extracranial/Intracranial Arterial Anastomosis (EC/IC Bypass Study): methodology and entry characteristics. **Stroke 16:**397–406, 1985
11. Executive Committee for the Asymptomatic Carotid Atherosclerosis Study: Endarterectomy for asymptomatic carotid artery stenosis. **JAMA 273:**1421–1428, 1995
12. Gates EA: New surgical procedures: can our patients benefit while we learn? **Am J Obstet Gynecol 176:**1293–1299, 1997
13. Haines SJ: Randomized clinical trials in the evaluation of surgical innovation. **J Neurosurg 51:**5–11, 1979
14. Kestle J, Milner R, Drake J: The Shunt Design Trial: variation in surgical experience did not influence shunt survival. **Pediatr Neurosurg 30:**283–287, 1999
15. Kestle JR, Drake JM, Cochrane DD, et al: Lack of benefit of endoscopic ventriculoperitoneal shunt insertion: a multicenter randomized trial. **J Neurosurg 98:**284–290, 2003
16. Lobato AC, Rodriguez-Lopez J, Diethrich EB: Learning curve for endovascular abdominal aortic aneurysm repair: evaluation of a 277-patient single-center experience. **J Endovasc Ther 9:** 262–268, 2002
17. McLeod RS: Issues in surgical randomized controlled trials. **World J Surg 23:**1210–1214, 1999
18. Moore WS, Vescera CL, Robertson JT, et al: Selection process for surgeons in the Asymptomatic Carotid Atherosclerosis Study. **Stroke 22:**1353–1357, 1991
19. Moore WS, Young B, Baker WH, et al: Surgical results: a justification of the surgeon selection process for the ACAS trial. The ACAS Investigators. **J Vasc Surg 23:**323–328, 1996
20. Neugebauer E, Troidl H, Spangenberger W, et al: Conventional versus laparoscopic cholecystectomy and the randomized controlled trial. Cholecystectomy Study Group. **Br J Surg 78:** 150–154, 1991
21. North American Symptomatic Carotid Endarterectomy Trial Collaborators: Beneficial effect of carotid endarterectomy in symptomatic patients with high-grade carotid stenosis. **N Engl J Med 325:**445–453, 1991
22. North American Symptomatic Carotid Endarterectomy Trial (NASCET) Steering Committee: North American Symptomatic Carotid Endarterectomy Trial. Methods, patient characteristics, and progress. **Stroke 22:**711–720, 1991
23. Ramsay CR, Grant AM, Wallace SA, et al: Assessment of the learning curve in health technologies. A systematic review. **Int J Technol Assess Health Care 16:**1095–1108, 2000
24. Ramsay CR, Wallace SA, Garthwaite PH, et al: Assessing the learning curve effect in health technologies. Lessons from the nonclinical literature. **Int J Technol Assess Health Care 18:** 1–10, 2002
25. Regan JJ, Yuan H, McAfee PC: Laparoscopic fusion of the lumbar spine: minimally invasive spine surgery. A prospective multicenter study evaluating open and laparoscopic lumbar fusion. **Spine 24:**402–411, 1999
26. Russell I: Evaluating new surgical procedures: needs collaboration between surgeons and trialist. **Brit Med J 311:**1243–1244, 1995
27. Schulze H: TULIP: transurethral ultrasound-guided laser-induced prostatectomy. **World J Urol 13:**94–97, 1995
28. Senagore AJ, Luchtefeld MA, Mackeigan JM: What is the learning curve for laparoscopic colectomy? **Am Surg 61:**681–685, 1995
29. Sequeira R, Weinbaum F, Satterfield J, et al: Credentialing physicians for new technology: the physician's learning curve must not harm the patient. **Am Surg 60:**821–823, 1994
30. Solomon MJ, McLeod RS: Should we be performing more randomized controlled trials evaluating surgical operations? **Surgery 118:**459–467, 1995
31. Stirrat GM, Farrow SC, Farndon J, et al: The challenge of evaluating surgical procedures. **Ann R Coll Surg Engl 74:**80–84, 1992

Manuscript received February 12, 2004.
Accepted in final form June 14, 2004.

Address reprint requests to: John R. W. Kestle, M.D., Division of Pediatric Neurosurgery, Primary Children's Medical Center, 100 North Medical Drive, Salt Lake City, Utah 84113. email: john.kestle@hsc.utah.edu.