OBSTETRICS

Use of the Learning Curve–Cumulative Summation test for quantitative and individualized assessment of competency of a surgical procedure in obstetrics and gynecology: fetoscopic laser ablation as a model

Ramesha Papanna, MD, MPH; David J. Biau, MD, PhD; Lovepreet K. Mann, MBBS;
Anthony Johnson, DO; Kenneth J. Moise Jr, MD

OBJECTIVE: We sought to determine the learning curve (LC) for fetoscopic laser photocoagulation (FLP) as a model for the evaluation of training in surgical procedures.

STUDY DESIGN: A retrospective review of consecutive case series of FLP from 2 centers with 3 operators (operator I [O-I], observer trained; operator II [O-II], hands-on trained; and operator III [O-III], clinical fellow) was performed and the LC-cumulative summation (CUSUM) test was plotted.

RESULTS: The acceptable and unacceptable success rates for at least 1 fetus survival after FLP were set at 82% and 70%, respectively, from a

systematic review. A total of 171 consecutive cases were performed by the 3 operators (O-I, 91; O-II, 49; and O-III, 31). From LC-CUSUM test O-I needed 60 procedures, O-II needed 20 procedures, and O-III needed 20 procedures to reach an acceptable performance rate for at least 1 survivor.

CONCLUSION: The LC-CUSUM test can be used to accurately assess the LC in a surgical procedure in obstetrics and gynecology. Hands-on trained operators exhibit a shorter LC.

Key words: formal training, laser photocoagulation, learning curve-cumulative summation, twin-twin transfusion syndrome

Proficiency for most surgical procedures is determined by 1 of 2 methods. A predetermined number of cases is prescribed–a technique that fails to take into account that all individuals may not achieve proficiency with a fixed sample size.[1] Alternatively, observation by a tutor is used to determine when competency has been reached–a method with considerable subjective bias.[2] Until recently, there has been very little interest in applying statistical methods to deter-

★ EDITORS' CHOICE ★

mine when a trainee becomes proficient at a particular surgical procedure. Biau et al[3] reviewed the surgical literature from 1991 through 2006 and noted 22 papers that used the cumulative summation (CUSUM) test to determine the learning curve (LC). CUSUM was first introduced into industry to assess the maintenance of proficiency once a process was "in control."[4] As an operator begins a

new procedure, by definition, the process is "out of control" until a level of steady proficiency is attained. In 2008, Biau et al[5] introduced a modification of the CUSUM method called LC-CUSUM that was specifically designed to determine when a level of proficiency has been attained. This new method objectively determines the LC for a procedure as to when the trainee has reached a predefined level of performance.[6]

The importance of establishing a LC is to reduce risks to patients of incompletely trained providers performing complex procedures. Additionally, evaluation of a LC can limit wastage of the available resources needed to train those who are already proficient. Thus both patient safety and cost containment are benefits that are realized. LC-CUSUM has been reported in studies for determining the LC for endoscopic retrograde cholangiopancreatographies.[5] Studies using this method in obstetrics and gynecology have been limited to embryo transfer for in vitro fertilization,[6] vitrification of the embryo,[7] and more recently

✎ For Editors' Commentary, see Table of Contents

becoming proficient at certain diagnostic ultrasound procedures.[8] The process of LC-CUSUM analysis has broader implications for training in surgical specialties, specifically in obstetrics and gynecology. Recently there has been increased emphasis by the Accreditation Council for Graduate Medical Education to assess the competency of a trainee, both qualitatively and quantitatively. Use of LC-CUSUM would help improve the efficiency in determining the training necessary to achieve competency for selected surgical procedures for residents and fellows.

The purpose of the present study was to determine the LC of fetoscopic-directed laser therapy in the treatment of twin-twin transfusion syndrome (TTTS) for 3 fetal interventionists who had undergone different methods of training. The CUSUM test was then employed to monitor the ongoing performance once competence had been reached.

MATERIALS AND METHODS

To establish acceptable and unacceptable success rates for fetoscopic laser photocoagulation (FLP), we performed a systematic review of published articles of consecutive case series (Appendix 1). A retrospective study was conducted from the review of our online database of cases that underwent FLP from June 2005 through September 2006 at University of North Carolina (UNC) at Chapel Hill and from November 2006 through August 2009 at the Texas Children's Fetal Center (TCFC) in Houston, TX. The study was approved by institutional review boards of both institutions. Fetal team members prospectively collected patient outcome data by direct patient contact and by contact with the referring physician.

There were 3 operators in the group. Operator I (O-I) was a board-certified maternal-fetal medicine specialist who was in active clinical and academic practice for >25 years with an equivalent number of years of experience in ultrasound-guided procedures. O-I was the first to learn the procedure in the group after an observation period in laser centers in Paris, France, and Leuven, Bel-

gium, for 1 month. Twelve live cases were observed between the 2 units, along with multiple hours of reviewing videotapes from previous cases. In addition, this operator visited a laser center in Tampa, FL, for a 2-day course on diagnosis and management of TTTS. The course included a series of lectures and videotape presentations on selective fetoscopic laser photocoagulation in the treatment of TTTS. O-I performed all procedures at UNC as the primary surgeon after an institutional review board approved performing the procedure. The patients were counseled about the experience of the surgeon during the consenting process for the procedure. Operator II (O-II) was a maternal-fetal medicine specialist in active clinical and academic practice for >25 years with an equivalent number of years of experience in ultrasound-guided procedures. O-II assisted O-I for FLP by providing ultrasound guidance during the procedure for the first 30 cases. O-II started performing the laser procedure at TCFC as the primary surgeon, initially performing only cases with a posterior placenta and then advancing to cases with an anterior placenta. O-I assisted in all cases during O-II's learning phase. Operator III (O-III) was a graduate from an American Board of Obstetrics and Gynecology–approved obstetrics and gynecology residency program and had performed >300 laparoscopic gynecologic procedures. O-III joined the fetal intervention team in July 2008 as a clinical fellow and observed 10 cases and assisted in the next 15 cases, before starting to perform FLP as the primary surgeon. In the beginning, O-III performed FLP on cases with posterior placentation (n = 15) and gradually transitioned to perform the procedure on cases with anterior placentation. O-I or O-II participated in all of O-III's cases and in the majority of cases both were present in the operating suite when the fellow was performing the procedure. The details and the evaluation of the TTTS patients and the procedure at our center have been previously described.[9,10]

The data for maternal demographics and for preoperative, intraoperative, and postoperative variables were extracted to

software (SPSS version 11.0; SPSS Inc, Chicago, IL). Comparison of parameters was performed using the χ^2 test for categorical variables and Fisher's exact test when an expected frequency was <5. Analysis of variance was performed for continuous variables. Kruskal-Wallis test was used to compare nonparametric variables among the 3 operators. A P value < .05 was considered as significant. The data needed for calculating LC-CUSUM and CUSUM included surgeon's identity, consecutive case series number for each surgeon, and number of surviving fetuses to birth for each pregnancy. These data were extracted to software (R-Software; R Foundation for Statistical Computing, Vienna, Austria).

The LC-CUSUM was developed to determine when a trainee has reached a predefined level of performance.[11] The LC-CUSUM sequentially tests the null hypothesis "performance is unacceptable" against the alternative "performance is acceptable." It computes a score from the successive outcomes, with successes yielding an increase in the score and failures yielding a decrease in the score. Once the score reaches a predefined limit (h), the test rejects the null hypothesis in favor of the alternative and performance is deemed acceptable.

Graphically the LC-CUSUM score is plotted on the y-axis against the successive procedures on the x-axis. As long as the score remains in the continuation region, namely between the x-axis and the decision limit h, performance cannot be considered as acceptable and monitoring continues. With accumulation of successes the score increases until it crosses the limit h where proficiency is declared. A particular feature of the LC-CUSUM is that it incorporates a holding barrier at 0 that cannot be crossed and the score St thus remains at 0 if the trainee accumulates numerous successive failures. In this way, the LC-CUSUM remains responsive at all times and if, for instance, poor performance resulted from poor technique, with improvement in skills the trainee will not have to compensate unnecessarily for all the accumulated failures and may be able to show acceptable performance in due course (Appendix 1). For the LC-CUSUM for at