Ultrasound-Guided Regional Anesthesia

How Much Practice Do Novices Require Before Achieving Competency in Ultrasound Needle Visualization Using a Cadaver Model

Michael J. Barrington, FANZCA,* Daniel M. Wong, FANZCA,* Ben Slater, FANZCA,*
Jason J. Ivanusic, PhD,† and Matthew Ovens‡

Background and Objectives: Ultrasound needle visualization is a fundamental skill required for competency in ultrasound-guided regional anesthesia. The primary objective of this study using a cadaver model was to quantify the number of procedures that novices need to perform before competency, using a predefined dynamic scoring system was achieved in ultrasound needle visualization skills.

Methods: Fifteen trainees, novices to ultrasound-guided anesthesia, performed 30 simulated sciatic nerve blocks in cadavers. After each procedure, a supervisor provided feedback regarding quality-compromising behaviors. Learning curves were constructed for each individual trainee by calculating cusum statistics. Trainees were categorized into those who were proficient, not proficient, and undetermined. A mathematical model predicted the number of procedures required before an acceptable success rate would be attained. Logistic regression was used to identify factors associated with success.

Results: There was wide variability in individual cusum curves. The mean number of trials required to achieve competency in this cohort was 28. Trainees were categorized as proficient (n = 6), not proficient (n = 5), and undetermined (n = 4). With each subsequent procedure, there was a significant increase in the likelihood of success for trainees categorized as not proficient ($P = 0.023$) or undetermined ($P = 0.024$) but not for trainees categorized as proficient ($P = 0.076$). Participants recruited later in the study had an increased likelihood of success ($P < 0.001$).

Conclusions: Trainees became competent in ultrasound needle visualization at a variable rate. This study estimates that novices would require approximately 28 supervised trials with feedback before competency in ultrasound needle visualization is achieved.

(*Reg Anesth Pain Med* 2012;37: 334–339)

Ultrasound needle visualization is a fundamental skill required for competency in ultrasound-guided regional anesthesia (UGRA). This skill requires a level of dexterity to achieve precise alignment of the needle and ultrasound beam. For many practitioners, acquiring this skill is a challenge, requiring practice and repetition. Ultrasound-guided regional an-

esthesia also requires competency in ultrasound machine use, scanning (transducer manipulation), and interpretation of sonographic images.[1] Novices to UGRA have been shown to make errors and quality-compromising behaviors (QCBs) in a clinical environment even after exposure to what seemed to be an appropriate volume of clinical material in a supervised environment.[2]

The environment where the complex motor skills required for clinical procedures are taught is important. In the operating room, production pressure, interruptions, unpredictability and other factors impair learning, and this is especially relevant for trainees learning regional anesthesia.[3] There is a trend for less reliance on exposure to sheer volume of cases in the operating room and more emphasis placed on a structured curriculum and preclinical procedural training.[4] Several models and trainers, including inanimate phantoms, virtual reality models and simulators, have been used to develop motor skills required for surgery and anesthesia.[5–10] Motor skills learned during preclinical procedural training can be transferred to the clinical environment. A bovine phantom model used to learn basic ultrasound skills showed that significant practice was required to achieve competency.[11] In addition to the trainer used, educational strategies, such as deliberate practice of component skills with feedback, may accelerate the rate of skill acquisition.[4] Cadavers are realistic in both sonoanatomy and tissue planes and can be used as a model to optimize trainee performance before clinical training. Despite these benefits, cadavers are expensive and have limited availability; therefore, their utility in teaching UGRA should be evaluated.

The primary objective of this study, using a cadaver model, was to quantify the number of procedures novices need to perform in a supervised, nonclinical environment before competency (using a predefined dynamic scoring system) was achieved in ultrasound needle visualization skills.

METHODS

This project was approved by the University of Melbourne, Human Research Ethics Committee. Nine frozen/thawed cadavers were used to simulate ultrasound needle visualization skills required for sciatic nerve blockade. Fifteen trainee anesthetists, novices to UGRA, were selected at random and recruited to perform the simulation in the presence of a supervisor (M.J.B.). Novices were defined as junior trainees who had performed less than 5 UGRA procedures. To provide the anatomic tissue necessary for 15 trainees, both legs were used from some cadavers. There were no trainees involved in the study observing other trainees perform procedures. This study was conducted during a 14-month period.

An M-Turbo ultrasound machine with a 38-mm, 13-6 MHz linear transducer (SonoSite, Inc, Bothell, Washington) covered in a protective plastic sheath, and a 150-mm, 20-gauge needle (B. Braun Stimuplex, Melsungen, Germany) were used for all procedures. The supervisor demonstrated to trainees individually

a mock in-plane, ultrasound-guided sciatic nerve block of the posterior thigh using an in-plane needle approach with the nerve in the short-axis. The cadavers were positioned prone and initial ergonomics optimized, including appropriate positioning of the ultrasound machine. The key requirements (nerve imaging, ultrasound needle visualization, injection) were demonstrated to all trainees before their first procedure. With respect to nerve imaging, these requirements comprised maintaining the target nerve two thirds of the distance across the screen in the short-axis. With respect to ultrasound needle visualization, these requirements comprised separating the needle skin entry point from the probe and using probe maneuvers (alignment, tilting, rotation, pressure, heel-toe) to provide a long-axis view of the needle and full visualization of the shaft and tip of the needle during the needle trajectory and injection. The procedure was considered complete when the needle tip was located at the 6- and 12-o'clock positions around the nerve, and 1 mL of saline was injected.

Before each procedure, if required, the supervisor clarified the location of the sciatic nerve and optimized image quality by adjusting standard machine settings (depth, gain, focus). During each procedure, no feedback was provided. After each procedure, the supervisor provided feedback in response to QCBs: advancement of needle while not visualized (QCB1), malposition of target nerve on screen (QCB2), poor probe handling (unintentional or ineffective probe movement; QCB3), awkward needle holding (QCB4), watching hands or needle instead of target (QCB5), failure to recognize intramuscular or maldistribution of injection (QCB6), intraneural injection (QCB7), fatigue (QCB8), failure to correlate sidedness of screen and probe (QCB9), and inappropriate needle insertion site (QCB10). Each trainee performed the procedure 30 times with a 15-minute break after the 10th and 20th procedures to reduce fatigue. Each procedure was performed approximately 1 cm remote to the previous one, so that the previous injectate would not distort the image plane.

The time taken to complete the procedure and videos of the procedures were recorded and then analyzed offline by an observer not present during the procedure and blinded to trainee identity and procedure number. A predefined dynamic scoring system was used to assess block performance under 2 categories: ultrasound needle visualization and probe steadiness. Ultrasound needle visualization was scored as: 0, needle was advanced with tip and whole shaft being fully visible; 1, needle was advanced with tip and part of shaft being visible; 2, needle was advanced once or twice without tip clearly visible; and 3, needle was advanced on 3 or more occasions without tip clearly visible. Transducer steadiness was scored as: 0, transducer steady and nerve imaged adequately in field of view; and 1, transducer unsteady, or nerve not adequately centered in the field of view, or multiple attempts at finding needle in plane. Also, captured on video for analyses was the visibility of needle tip during injection (0, visible; 1, tip not visible).

Learning curves were constructed for each individual trainee by calculating cusum statistics using the following parameters: probability of type I error (α) = 0.05, probability of type II error (β) = 0.2, acceptable failure rate = 10%, and unacceptable failure rate = 20%.[12] The sample size for the minimum number of trials that each trainee was required to perform was estimated to be 25 to 29 (Appendix 1). Because this study was an observational trial, no calculation was required to determine the number of trainees that needed to be recruited. Nonetheless, 15 were chosen because we considered it likely that this number would represent a reasonable cross section of trainees.

The definition of success used to construct the cusum curve was a dynamic score of less than 3. Trainees were categorized as proficient if their cusum curve crossed the lower limit line h_0 from above and not proficient if the cusum curve crossed the upper decision limit h_1 from below. When the cusum curve was located between the upper and lower decision limit, trainee proficiency was categorized as undetermined. Mean learning curves were constructed using median cusum statistics for the 3 proficiency categories. Depth of the target nerve (measured from the 12-o'clock position to the skin), procedural time, trial number, participant number, and cadaver code were used as covariates in a logistic regression model to identify factors associated with
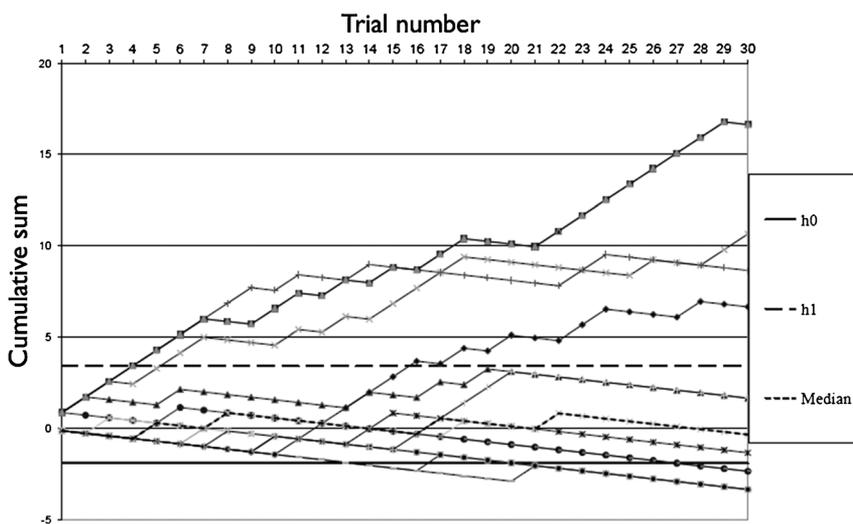
FIGURE 1. The individual cusum curves of 15 trainees. Upward and downward deflection indicates failure and success, respectively. The cusum curves of 6 trainees crossed lower decision limit, h_0 from above, indicating competency was achieved. The cusum curves of 5 trainees crossed the upper decision limit, h_1 from below, indicating competency was not achieved. The cusum curves of 4 trainees' lines were located between upper and lower decision limits, indicating no statistical inference could be made about their competency. Competency designated at trial number 30.

335
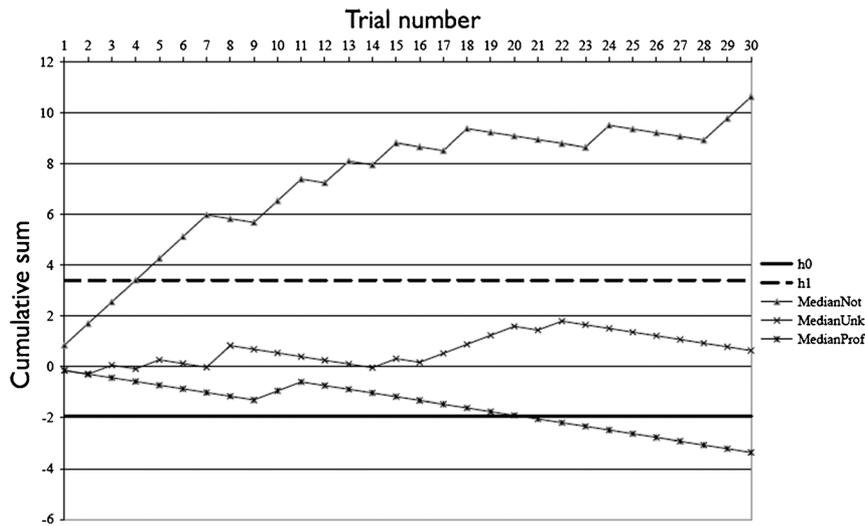
FIGURE 2. MedianNot refers to median cusum statistics for trainees who did not achieve competency. MedianUnk indicates median cusum statistics for trainees for whom no statistical inference could be made about their competency. MedianProf indicates median cusum statistics for trainees who achieved competency. Competency designated at trial number 30. Lower (h_0) and upper (h_1) decision limits are indicated.

success.[13] The Spearman ρ correlation was used to evaluate a correlation between the ultrasound needle visualization and visibility of needle tip at injection scores. $P < 0.05$ was considered statistically significant.

To predict the number of procedures that would be required before an acceptable success rate would be attained, data (successes and failures) were used to construct a Bush and Mosteller mathematical learning model aiming at a probability of success equal to 90% while assuming as acceptable a 20% probability of failure (Appendix 2).[14] An MS Excel electronic spreadsheet (Microsoft, Bellevue, Washington) was used to construct cusum curves and Bush and Mosteller's mathematical learning model. All other analyses were performed using SPSS 17 (IBM North America, New York, New York). Except when otherwise indicated, data are summarized by medians (quartiles), and success rates are expressed in a 0 through 1 scale.

RESULTS

The cusum learning curves for the 15 individual trainees are demonstrated in Figure 1. There was wide variability in the individual curves. The median cusum statistics of trainees, categorized into those who were proficient (n = 6), those who were not proficient (n = 5), and those where no statistical inference could be made (undetermined, n = 4), are summarized in Figure 2. The mean number of trials required to achieve competency in this cohort was estimated at 28 (Fig. 3). The distribution of QCBs that provided the basis for feedback after each trial is summarized in Figure 4.
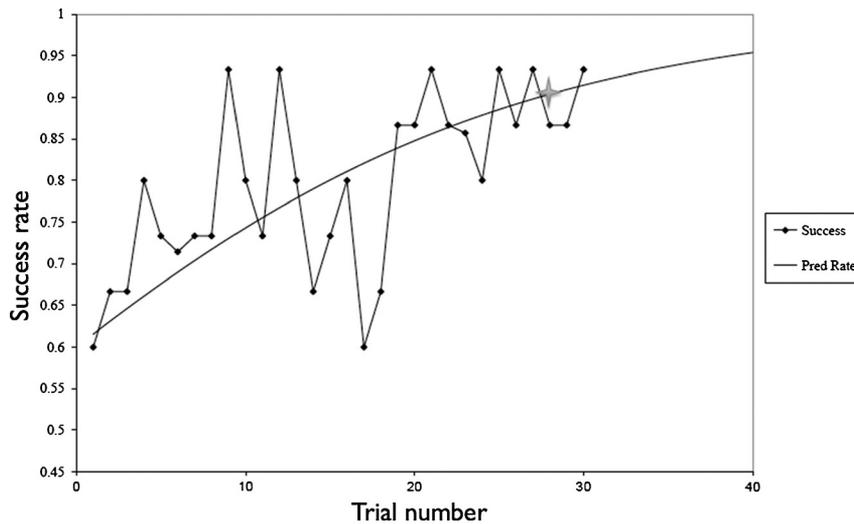
FIGURE 3. The observed success rates (Success) and the predicted success rate (Pred Rate) rates at each trial, as predicted by Bush and Mosteller's learning model, using data from all trainees are shown. The model predicts that a 90% success rate would be attained after an average of 28 trials (star indicates where curve crosses 90% mark).

336
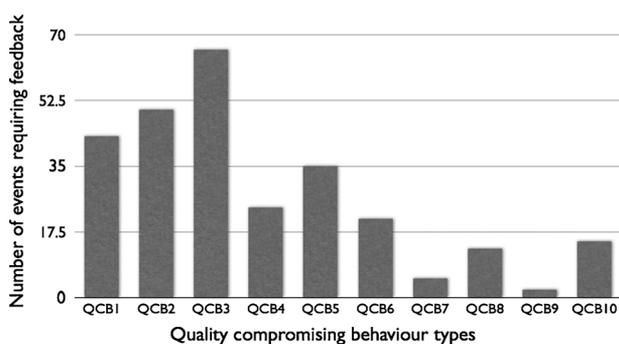
FIGURE 4. Distribution of quality-compromising behaviors (QCBs) that provided the basis for feedback. In total, there were 274 QCBs (from 450 trials) identified by the supervisor. QCB 1 (16%) indicates advancement of needle while not visualized; QCB 2 (18%), malposition of target nerve on screen; QCB 3 (24%), poor probe handling; QCB 4 (9%), awkward needle holding; QCB 5 (13%), watching hands or needle instead of target; QCB 6 (8%), failure to recognize intramuscular or maldistribution of injection; QCB 7 (2%), intraneural injection; QCB 8 (5%), fatigue; QCB 9 (1%), failure to correlate sidedness of screen and probe; QCB 10 (5%), inappropriate needle insertion site.

After each subsequent trial, there was a statistically significant increase in the likelihood of success for trainees categorized as not proficient ($P = 0.023$) or undetermined ($P = 0.024$), but no statistically significant increase for trainees categorized as proficient ($P = 0.076$). When expressed as odds ratios, the estimates of likelihood of success were 1.046 (95% confidence interval [CI], 1.006–1.088), 1.098 (95% CI, 1.013–1.191), and 1.059 (95% CI, 0.994–1.128) for the not proficient, undetermined, and proficient categories, respectively. Increased procedural duration was associated with a higher probability of failure ($P < 0.001$). There was no statistically significant reduction in procedural time with trial number. The mean depth of nerve was 1.41 cm (range, 0.98–2.09 cm), and there was no significant association between this parameter or cadaver code and likelihood of success. There was a significant association between the ultrasound needle visualization and visibility of needle tip at injection scores (Spearman ρ correlation of 0.335, $P < 0.001$). Participant number had a statistically significant effect ($P < 0.001$) on the trial outcome, with an estimated odds ratio of 1.138 (95% CI, 1.075–1.206). That is, participants recruited later in the study had an increased likelihood of success.

DISCUSSION

The results of this training model using cadavers indicate that trainees who are novices to UGRA require approximately 28 supervised trials with feedback to achieve competency in ultrasound needle visualization skills. There was wide variability in the learning curves of individual trainees; however, this is consistent with a previous study.[11] The likelihood of success increased significantly with each subsequent trial, both for trainees who were categorized as not proficient (crossed the upper decision limit h_1 from below) or where proficiency was undetermined (between the decision limits). In clinical practice, recognizing that a trainee is not proficient is important, indicating that further supervised practice is warranted. The benefit of each additional trial in increased likelihood of success was small (approximately 5%); this is consistent with clinical observations that UGRA is a demanding skill to learn and it provides evidence that improved performance took place using this cadaver model of training. In contrast, trainees categorized as proficient (their cusum curve crossed the lower limit line h_0 from above) did not benefit from further practice, and in clinical practice, reducing their supervision and assigning more challenging tasks may be more appropriate.

The association between increased procedural duration and higher probability of failure is not surprising as an increased level of procedural expertise is typically associated with reduced procedural duration. The depth of the target nerve and the cadaver could potentially have been confounders influencing success, but this was not shown in our results. The relatively shallow depth of the nerve may have contributed to this parameter not being a predictor of success. Participants recruited later in the study had a greater likelihood of success, indicating that perhaps they received clearer instructions as the supervisor became more proficient in delivering them. This further highlights the importance of adequate supervision and training. An additional potential confounder is that trainees recruited later in the study had their learning contaminated by increased exposure to ultrasound in clinical practice other than regional anesthesia. With these potential sources of bias in mind, readers should note that the findings of this study may not necessarily translate directly to their own clinical practice and that different training methods may yield different results. The association between ultrasound needle visualization and visibility of needle tip at injection reinforces the importance of identifying the trainee with poor dynamic needle visualization skills because the needle tip at the point of injection is also likely to be poorly visualized.

A study by de Oliveira Filho et al[11] using a bovine phantom constructed learning curves and mathematical learning models for basic skills required for URGA: optimizing needle–ultrasound beam alignment and reaching a target inside a phantom. The results of that study indicated that 37 and 109 trials would be required to achieve proficiency in optimizing needle–ultrasound beam alignment (experiment 1) and reaching a target inside a phantom (experiment 2), respectively. We used similar analytical methods in this current study on a cadaver model simulating a sciatic nerve block. Initially, our study was designed to include perineural spread as one component used to determine success, and therefore, our procedure would have been similar to experiments 1 and 2 combined. However, the cadaver model frequently demonstrated the inappropriate spread of injectate (eg, intramuscular) despite ideal needle tip position, and therefore, we omitted this from the score determining success. This unexpected spread of injectate may have occurred due to postmortem changes in tissue and represent a limitation of using cadavers for training. However, this can also occur in human clinical practice, and the ability of a clinician to recognize inappropriate spread of local anesthetic and then correct needle position is important.

In the study by de Oliveira Filho et al,[15] no instruction or feedback was provided to the learner between trials; this educational approach has been termed *discovery learning*. This current study used deliberate practice, an educational strategy associated with the level of expertise acquired in a wide range of domains including music, chess, and sport.[4] Deliberate practice includes providing the trainee with feedback. The 3 most common QCBs requiring feedback were poor transducer handling, malposition of the target nerve on the screen, and advancement of needle while not visualized (Fig. 4). Advancing the needle out of plane (QCB1) was identified in 16% of trials, in contrast to a clinical study by Sites et al[2] where this error was the most common error identified (43%). This difference may potentially be due to the feedback trainees in this current study received after each individual trial. In addition, the relatively shallow needle trajectory and the cadaveric tissue itself

resulted in the needle being relatively easy to visualize, reducing the incidence of QCB1.

Potentially, gel phantoms can be used to develop the motor skills required for UGRA. Gel phantoms are reusable, and compared with cadavers, they are cost-effective and readily available. The cadaveric model in this study allowed a more realistic simulation by providing trainees with the opportunity to learn ultrasound needle visualization skills required for in-plane techniques with a real sciatic nerve and surrounding tissues. Moreover, the cadaveric tissue provided realistic tactile feedback as the needle was advanced. During the study, it became apparent that trainees appreciated the fidelity of the model and the opportunity to perform 30 sequential procedures. In a recent review describing the use of phantoms to practice needle guidance skills, it was noted that cadavers retain much of the textural feel of live human tissue and are nearly as echogenic.[16] Despite this, the restricted availability of cadavers and the requirement for training to take place in a location remote from the operating room reduce their utility as a training model.

Ultrasound needle visualization is an important generic skill required for in-plane UGRA. Although we used a sciatic nerve block model, the skills developed using this model are skills generic to any in-plane approach. Sciatic nerve was chosen so that the trainee could practice a similar but not identical procedure over the length of the nerve. An alternative would have been for the novice to practice at the exact same anatomic location. However, this would have reduced the learning potential of the repetitive tasks and potentially reduced the fidelity of the model. The skills acquired by trainees during this study should be transferable to the clinical environment; however, that was not a part of this project. Skills acquired, using an inanimate airway endoscopy trainer, and epidural simulator were transferable to the clinical environment.[5,6] It should also be emphasized that the skill of maintaining the needle in the plane of the ultrasound is only one key skill required to obtain overall proficiency in UGRA. Individual skill sets are required for ultrasound machine use and interpretation of sonograms. By acquiring these skills separately, the trainee may be able to more easily integrate the skills in the operating room environment.

Cusum statistical techniques were initially developed as an industrial quality control tool and have been used to monitor surgical performance[17–19] and document learning curves and assess competency for a wide range of procedures.[12,20–24] Cusum curves are easy to construct and provide a graphical format to document performance. The potential limitations and challenges in using cusum have been reviewed and include appropriate choice of input parameters: acceptable failure rate, unacceptable failure rate, probability of type 1 and type 2 errors, and definition of success.[12,24] The acceptable failure rate used in this current study was 10% based on a training study in a clinical environment that used cusum statistics.[21]

In conclusion, trainees became competent in ultrasound needle visualization at a variable rate. This study estimates that novices from a similar cohort would require approximately 28 supervised trials before competency in ultrasound needle visualization is achieved when using a cadaver model to simulate sciatic nerve block with supervision and feedback. There was evidence that improved performance took place using this training model.

ACKNOWLEDGMENT

REFERENCES

1. Sites BD, Chan VW, Neal JM, et al. The American Society of Regional Anesthesia and Pain Medicine and the European Society of Regional Anaesthesia and Pain Therapy joint committee recommendations for education and training in ultrasound-guided regional anesthesia. *Reg Anesth Pain Med*. 2010;35:S74–S80.

2. Sites BD, Spence BC, Gallagher JD, Wiley CW, Bertrand ML, Blike GT. Characterizing novice behavior associated with learning ultrasound-guided peripheral regional anesthesia. *Reg Anesth Pain Med*. 2007;32:107–115.

3. Jones RW, Morris RW. Facilitating learning in the operating theatre and intensive care unit. *Anaesth Intensive Care*. 2006;34:758–764.

4. Reznick RK, MacRae H. Teaching surgical skills—changes in the wind. *N Engl J Med*. 2006;355:2664–2669.

5. Martin KM, Larsen PD, Segal R, Marsland CP. Effective nonanatomical endoscopy training produces clinical airway endoscopy proficiency. *Anesth Analg*. 2004;99:938–944, table of contents.

6. Friedman Z, Siddiqui N, Katznelson R, Devito I, Bould MD, Naik V. Clinical impact of epidural anesthesia simulation on short- and long-term learning curve: high- versus low-fidelity model training. *Reg Anesth Pain Med*. 2009;34:229–232.

7. Gallagher AG, Cates CU. Approval of virtual reality training for carotid stenting: what this means for procedural-based medicine. *JAMA*. 2004;292:3024–3026.

8. Matsumoto ED, Hamstra SJ, Radomski SB, Cusimano MD. The effect of bench model fidelity on endourological skills: a randomized controlled study. *J Urol*. 2002;167:1243–1247.

9. Wanzel KR, Matsumoto ED, Hamstra SJ, Anastakis DJ. Teaching technical skills: training on a simple, inexpensive, and portable model. *Plast Reconstr Surg*. 2002;109:258–263.

10. Sedlack RE, Kolars JC. Computer simulator training enhances the competency of gastroenterology fellows at colonoscopy: results of a pilot study. *Am J Gastroenterol*. 2004;99:33–37.

11. de Oliveira Filho GR, Helayel PE, da Conceicao DB, Garzel IS, Pavei P, Ceccon MS. Learning curves and mathematical models for interventional ultrasound basic skills. *Anesth Analg*. 2008;106: 568–573, table of contents.

12. Kestin IG. A statistical approach to measuring the competence of anaesthetic trainees at practical procedures. *Br J Anaesth*. 1995;75:805–809.

13. Field AP. *Discovering Statistics Using SPSS*. 3rd ed. London, UK: SAGE Publications; 2009.

14. Bush RR, Mosteller F. A mathematical model for simple learning. *Psychol Rev*. 1951;58:313–323.

15. Vereijken B, Whiting HT. In defence of discovery learning. *Can J Sport Sci*. 1990;15:99–106.

16. Hocking G, Hebard S, Mitchell CH. A review of the benefits and pitfalls of phantoms in ultrasound-guided regional anesthesia. *Reg Anesth Pain Med*. 2011;36:162–170.

17. Grunkemeier GL, Wu YX, Furnary AP. Cumulative sum techniques for assessing surgical results. *Ann Thorac Surg*. 2003;76:663–667.

18. Biau DJ, Resche-Rigon M, Godiris-Petit G, Nizard RS, Porcher R. Quality control of surgical and interventional procedures: a review of the CUSUM. *Qual Saf Health Care*. 2007;16:203–207.

19. Yap CH, Colson ME, Watters DA. Cumulative sum techniques for surgeons: a brief review. *ANZ J Surg*. 2007;77:583–586.

20. East JM, Valentine CS, Kanchev E, Blake GO. Sentinel lymph node biopsy for breast cancer using methylene blue dye manifests a short learning curve among experienced surgeons: a prospective tabular cumulative sum (CUSUM) analysis. *BMC Surg*. 2009;9:2.

21. Naik VN, Devito I, Halpern SH. Cusum analysis is a useful tool to assess resident proficiency at insertion of labour epidurals. *Can J Anaesth*. 2003;50:694–698.

22. Sibanda T, Sibanda N. The CUSUM chart method as a tool for continuous monitoring of clinical outcomes using routinely collected data. *BMC Med Res Methodol*. 2007;7:46.

23. Schuepfer G, Johr M. Psoas compartment block (PCB) in children: part II—generation of an institutional learning curve with a new technique. *Paediatr Anaesth*. 2005;15:465–469.

24. Bolsin S, Colson M. The use of the cusum technique in the assessment of trainee competence in new procedures. *Int J Qual Health Care*. 2000;12:433–438.

APPENDIX 1

Constructing a cusum curve requires values for p_0, the acceptable failure rate; p_1, the unacceptable failure rate; α, the probability of type 1 error; and β, the probability of type 2 error. The cusum formulae allow calculation of the parameters s; h_1, the upper decision limit; and h_0, the lower decision limit: $a = \ln[(1 - \beta)/\alpha]$; $b = \ln[(1 - \alpha)/\beta]$; $P = \ln(p_1/p_0)$; $Q = \ln[(1 - p_0)/(1 - p_1)]$; $s = Q/(P + Q)$; $h_0 = -b/(P + Q)$; $h_1 = a/(P + Q)$. The cusum chart starts at 0. For each successive success, s is subtracted from the cusum score. For each failure, $(1 - s)$ is added to the score. At any point on the curve (C_n), the number of successes in successive trials needed to reach h_0 can be calculated as $(C_n - h_0)/s$. The average size of samples with failure rates $= p_0$ is calculated by $[(h_0 (1 - \alpha) - \alpha h_1)/(s - p_0)]$. For samples with failure rates $= p_1$, the average size is estimated by $[(h_1 (1 - \beta) - \beta h_0)/(p_1 - s)]$.

APPENDIX 2

The mean probability of success at trial $n(V_{1,n})$ can be estimated from Bush and Mosteller's learning model: $V_{1,n} \cong V_{1,0} /[V_{1,0} + (1 - V_{1,0})e^{-(\pi_1 - \pi_2)(1 - \alpha_1)n}]$, where $V_{1,0} = 0.5$ or the average of the successes rates at the initial trials; $\alpha_1 =$ the slope parameter; $n =$ trial number. The parameter α_1 is calculated by the formula: $\alpha_1 = 1 - [(\pi_1 - V_{1,0})/(N\pi_1 - T_1)]$, where $N =$ number of trials, $T_1 =$ mean number of successes at the $N - 1$ trials.