

# The Effectiveness of Video Technology As an Adjunct to Teach and Evaluate Epidural Anesthesia Performance Skills

David J. Birnbach, M.D.,\* Alan C. Santos, M.D., M.P.H.,† Remi A. Bourlier, M.D.,‡ Warner E. Meadows, M.D.,§ Sanjay Datta, M.D., F.F.A.R.C.S.,|| Deborah J. Stein, M.D.,# Maxine M. Kuroda, Ph.D., M.P.H.,\*\* Daniel M. Thys, M.D.†

**Background:** Although video review has been used in teaching, it has not been reported for use as an adjunct to teaching anesthesiology residents. The purpose of the prospective, randomized, blinded study was to determine whether teaching with video review improves epidural anesthesia skills of anesthesiology residents.

**Methods:** Twenty-two second-year (CA-2) anesthesiology residents beginning their first obstetric anesthesia rotation were assigned to video or nonvideo groups. All residents were filmed daily as they placed epidural analgesia. Residents assigned to the video group reviewed their tapes twice a week with an attending anesthesiologist, whereas residents assigned to the nonvideo group never saw their films. Four experienced attending anesthesiologists independently judged videotapes taken on days 1, 15, and 30 and scored the residents for "overall" skill (range of summed overall grades, 0–40), as well as on 13 predetermined criteria.

**Results:** As determined by kappa coefficients, interrater reliability was high among the judges ( $k = 0.7-0.8$ ). Residents in the video group improved to a greater degree than residents in the nonvideo group. On day 1, the median overall grades for the video and nonvideo groups were 21 and 12, respectively. By day 15, the corresponding grades had increased to 32 and 24, respectively ( $P < 0.01$ ). However, overall median grades continued to improve between days 15 and 30 in the video group only ( $P < 0.01$ ).

**Conclusions:** Review of resident videotapes resulted in greater improvement in overall and predetermined performance criteria. In addition, video review was helpful in identifying skills

that were inadequately learned, thus allowing for specific teaching in those areas.

CLINICAL teaching of obstetric anesthesia can provoke anxiety for all parties involved. Parturients requesting regional analgesia for pain relief during labor are often in intense pain and want the procedure performed expeditiously without the pauses that may be necessary for maximum resident education. Women in labor, in contrast to surgical patients, have not been given any sedative or hypnotic drugs and thus are more anxious and acutely aware of conversations regarding the performance of regional anesthesia. In addition, residents in obstetric anesthesia may feel rushed to initiate pain relief and may not pay sufficient attention to developing meticulous technique. Also, supervisory staff may feel reticent to offer suggestions for improvement at the point of care, particularly if the patient is already anxious and in pain.

The American Board of Anesthesiology has identified technical facility, medical judgment, and scholarship as the criteria on which competence is based.<sup>1</sup> Technical facility when performing obstetric anesthesia blocks may be difficult to teach and to assess. Video filming has been used extensively by professional athletes and in several areas of medical education to teach particular techniques and improve performance.<sup>2-5</sup> There may be several advantages to video filming of residents as an adjunct to conventional teaching of regional anesthesia techniques in obstetric patients. Residents and supervisory attending physicians can review and critique performance when not directly involved in patient care because the technique has been captured on film. Teaching can also be tailored to the individual strengths and weaknesses of the resident. A video library of the most common or unusual missteps can be created as an instructional device. Program directors and staff can evaluate and follow the progress of residents. Most important, video filming may decrease the amount of patient anxiety generated by numerous heuristic discussions during the performance of the block at the point of care.

The purpose of the current study was to assess the effectiveness of video filming as an adjunct for teaching obstetric regional analgesia techniques to residents.

This article is accompanied by an Editorial View. Please see: Gaba DM: Two examples of how to evaluate the impact of new approaches to teaching. ANESTHESIOLOGY 2002; 96:1-2.

\* Associate Professor, Departments of Anesthesiology and Obstetrics and Gynecology, # Assistant Professor, † Professor, Department of Anesthesiology, College of Physicians and Surgeons of Columbia University. ‡ Staff Anesthesiologist, Clinique de la Muette, Paris, France. § Attending Anesthesiologist, Beth Israel Medical Center, New York, New York. || Professor of Anesthesiology, Harvard Medical School. \*\* Epidemiologist, Children's Memorial Hospital, Northwestern University School of Medicine, Chicago, Illinois.

Received from the Department of Anesthesiology, St. Luke's-Roosevelt Hospital Center, College of Physicians and Surgeons of Columbia University, New York, New York, and the Department of Anesthesia, Perioperative and Pain Medicine, Brigham and Women's Hospital, Harvard Medical School, Boston, Massachusetts. Submitted for publication October 5, 2000. Accepted for publication June 8, 2001. Supported in part by a grant from the Society for Education in Anesthesia, Richmond, Virginia (to Dr. Birnbach). Presented in part at the annual meeting of the Society for Education in Anesthesia, San Antonio, Texas, May 29-31, 1997, and the annual meeting of the Society for Obstetric Anesthesia and Perinatology, Montreal, Quebec, Canada, June 1-4, 2000.

Address reprint requests to Dr. Birnbach: Director of Obstetric Anesthesiology, St. Luke's-Roosevelt Hospital Center, 1000 Tenth Avenue, New York, New York 10019. Address electronic mail to: djb2@columbia.edu. Individual article reprints may be purchased from the Journal Web site, www.anesthesiology.org.

## Materials and Methods

The study was approved by the Institutional Review Board (St. Luke's-Roosevelt Hospital Center, College of Physicians and Surgeons of Columbia University, New York, New York), and written informed consent was obtained from laboring women and residents to be filmed in the study. Twenty-two CA-2 anesthesiology residents beginning their 1-month rotation on the labor and delivery ward were assigned in random order to one of two groups: video review (VR) (n = 11) or nonvideo review (NVR) (n = 11).

All residents were directly supervised by an attending anesthesiologist who was unaware of the resident's group assignment and who was free to instruct as necessary. An obstetric anesthesiology fellow videotaped all residents daily as they placed epidural analgesia, and ensured that the faces of the residents were not recorded on film. A high-8 video camera with 24× zoom capability was used for filming. Resolution was of sufficient quality to allow clear visualization of the markings on the epidural catheter. Residents assigned to the VR group reviewed their tapes twice a week with an attending obstetric anesthesiologist and were asked to identify the technical errors that they had made (self-assessment); then the attending would further review the video with the resident. Although residents assigned to the NVR group did not view their video tapes, they were given technique teaching sessions with the same anesthesiologist. All residents, regardless of group, attended a daily didactic session on obstetric anesthesia.

The Inter-Hospital Group for Anesthesia Education has proposed 35 criteria, some of which can potentially be evaluated by video analysis.<sup>6</sup> To identify criteria that could be reliably evaluated in laboring women, the videotapes of six residents administering epidural analgesia were studied by four independent anesthesiologists before the launch of the study. Thirteen of the 35 criteria proposed by the Inter-Hospital Group were deemed by all four reviewers to be clinically significant to the performance of an epidural procedure in a laboring woman. The kappa coefficients for each of the 13 predetermined skills indicated excellent to very good interrater reliability among the judges (table 1). All residents had knowledge of the 13 specific criteria being evaluated.

For the current study, four experienced obstetric anesthesiologists served as judges and independently viewed the videotapes that were taken of the residents on days 1, 15, and 30 of their rotation. These time intervals were chosen to coincide with the beginning, middle, and end of their 1-month subspecialty rotation in obstetric anesthesia. On each day, each judge assigned scores of 0 (a major error), 1 (a minor error), or 2 (no error) in each specific skill criterion (thus, the range of possible scores for each skill was 0–8). This grading system was chosen so that, when the individual criterion

**Table 1. Thirteen Criteria the Judges Were Instructed to Grade**

|                               |  |
|-------------------------------|--|
| Proper positioning            | Was the patient seated evenly ( <i>i.e.</i> , not on an angle) on the bed?   |
|                               | Was the patient asked to "round" her back?   |
|                               | Was the resident's position optimal for placement of the epidural?   |
|                               | Was the bed height appropriate?  |
|                               | After completion of the epidural, was uterine displacement attained?   |
| Adhering to aseptic technique | Were gloves put on sterilely?  |
|                               | Was the sterile field contaminated at any point?   |
|                               | Was the kit opened aseptically?  |
|                               | Was the patient's back circumferentially prepared?   |
|                               | Were three separate sponge sticks applied?   |
|                               | Was sterile technique maintained during catheter insertion and taping?   |
| No spillage                   | Was the povidone iodine solution poured without contaminating the epidural kit?  |
| Povidone iodine               | Was povidone iodine solution allowed to dry?   |
|                               | Was adequate time allowed between sponge stick applications?   |
| Draping                       | Was the drape placed without contaminating the block site, gloves, or epidural kit?                                      |
| Needle control                | Was the needle advanced carefully and steadily?  |
| Midline location*             | Was the needle placed at the correct interspace ( <i>i.e.</i> , below the L2–L3 interspace) and in the midline position? |
| Cephalad direction            | Was the needle incorrectly inserted in a caudad direction?   |
| Air amount                    | Was an excessive amount (> 5 ml) of air injected into the epidural needle?   |
| Catheter insertion            | Was the catheter advanced carefully?   |
|                               | Was the length of catheter appropriate ( <i>i.e.</i> , approximately 5 cm into the epidural space)?                      |
|                               | Was the catheter withdrawn to the proper length, if necessary?   |
| Needle withdrawal             | Was the needle withdrawn carefully and steadily without displacing the catheter?   |
|                               | If the catheter could not be threaded, was the needle and catheter withdrawn together?                                   |
| Catheter securing             | Was the catheter taped appropriately?  |
|                               | Did the resident avoid the patient's hair during taping?   |
|                               | Was the catheter aspirated?  |
|                               | Was the epidural catheter secured without kinking or displacement?   |
| Gentle treatment              | Was the patient told that the disinfectant solution was about to be applied?   |
|                               | Was the patient told of the impending needle stick prior to administration of local anesthetic?                          |
|                               | Was local anesthetic administered in an appropriate manner?  |
|                               | Was the patient informed of the possibility of paresthesia at the time of catheter insertion?                            |
|                               | After taping, was the patient placed in a comfortable, noninjurious position?  |

\* Residents at this institution are taught to use the midline approach.

scores were summed over the 13 criteria (possible range of summed scores, 0–26), higher summed criterion scores would indicate better technique.

In addition to the criterion scores, each judge assigned an “overall” grade to each resident on days 1, 15, and 30. These grades of 0–10 (possible range of summed overall grades, 0–40) were based on the judge’s subjective evaluation of the resident’s general performance and thus were obtained independent of the scores summed over the 13 criteria. Overall grades were obtained so that they could be correlated with the summed scores and could be used to assess whether a few poorly performed skills in certain criteria could lower the summed scores of a resident who was performing well overall.

Because an increase of at least three points was deemed necessary to detect notable improvement for any skill, only residents who could have achieved a three-point improvement in their scores were considered for this analysis. Thus, residents who already scored high on a skill (e.g., were given scores of 2 from each of the judges for a skill) would not be able to improve any further on that criterion and could not be judged for improvement.

#### Statistical Analysis

For each criterion score, interrater reliability among judges was assessed by the kappa coefficient. This coefficient, which indicates agreement among the judges after chance agreement is eliminated, was calculated for each pair of judges and then averaged over the six possible pairs. Using the denominators of the individual kappa values as “weights,” a cumulative kappa coefficient incorporating the 13 techniques was also calculated following the procedures of Fleiss.<sup>7</sup>

Overall grades and summed criterion scores were dichotomized at the median because their distributions were nonnormal. Thus, the association between overall grade and summed score was analyzed by chi-square test for days 1, 15, and 30. Because of the nonnormal distribution of the scores on the 13 criteria, the Mann-Whitney U test was used to analyze differences between the VR and NVR groups. The Wilcoxon matched-pairs, signed rank test was used to test differences among days 1, 15, and 30 within each of the groups. *P* values less than 0.05 were considered to be statistically significant. All analyses were performed with the Statistical Package for the Social Sciences (version 5.02 for Windows; SPSS, Chicago, IL).

## Results

All instances of epidural analgesia functioned well initially. There was no significant difference in the number of epidural procedures performed by residents in the VR and NVR groups (mean  $\pm$  SD: 73  $\pm$  6 and 71  $\pm$  4,

**Table 2. Kappa Coefficients\* for Each of the Thirteen Skills Selected from the Inter Hospital Study Group for Anesthesia Education**

|                            | Kappa |
|----------------------------|-------|
| Proper positioning         | 0.8   |
| Aseptic technique          | 0.7   |
| No spillage                | 0.8   |
| Povidone iodine dried      | 0.8   |
| Draping                    | 0.7   |
| Needle control             | 0.8   |
| Midline location           | 0.8   |
| Cephalad direction         | 0.7   |
| Air amount                 | 0.8   |
| Catheter insertion         | 0.8   |
| Needle withdrawal          | 0.7   |
| Catheter securing          | 0.7   |
| Patient treated gently     | 0.8   |
| Cumulative criteria score† | 0.8   |

\* Summed for days 1, 15, and 30. † Summed for the 13 skills.

respectively). A total of six unintentional dural punctures (“wet taps”) occurred during the study, one of these by a resident in the VR group during the first week of the rotation. The remaining five occurred in the NVR group during the third and fourth weeks of rotation.

There was good to excellent agreement among the four judges for each criterion evaluated because the kappa coefficients ranged from 0.7 to 0.8 for the 13 criteria (table 2). The kappa coefficient for the cumulative criterion scores was also excellent, being 0.8.

Interrater reliability was also high for the overall grades (kappa coefficient = 0.8). Overall grades were highly associated with summed scores for the 13 study criteria (chi-square test, *P* < 0.0001 for all 3 days). Compared to their day-1 grades, the median overall grades of residents in both groups were significantly higher by day 15. However, overall grades continued to improve only in the VR group, achieving a score of 36 by day 30 (*P* < 0.01) (table 3). Likewise, median total scores showed a statistically significant change between days 1 and 15 for both VR and NVR groups but only for the VR group between days 15 and 30 (table 4).

The percentage of residents with at least a three-point improvement in skills between days 1 and 30 is given in table 5. By day 30, almost all residents in the VR group

**Table 3. Median Overall Grades (Range) for VR and NVR Groups on Days 1, 15, and 30**

|         | VR         | NVR        | <i>P</i> Value* |
|---------|------------|------------|-----------------|
| Day 1   | 21 (4–35)  | 12 (8–34)  | NS              |
| Day 15† | 32 (14–38) | 24 (16–38) | 0.018           |
| Day 30‡ | 36 (26–40) | 24 (19–38) | 0.001           |

Range of summed overall grades, 0–40 on each evaluable day.

\* Mann-Whitney U test comparing video review (VR) and nonvideo review (NVR) groups on days 1, 15, and 30. † *P* < 0.01 for both VR and NVR groups comparing days 1 and 15 by Wilcoxon test. ‡ *P* < 0.01 for VR group only comparing days 15 and 30 by Wilcoxon test.

NS = not significant.

**Table 4. Median Total Scores (Range) for VR and NVR Groups on Days 1, 15, and 30**

|         | VR           | NVR         | P Value* |
|---------|--------------|-------------|----------|
| Day 1   | 54 (0–99)    | 46 (1–103)  | NS       |
| Day 15† | 100 (33–104) | 81 (34–104) | 0.09     |
| Day 30‡ | 104 (75–104) | 82 (46–104) | 0.002    |

Range of summed criteria scores, 0–104 for each evaluable day.

\* Mann–Whitney U test comparing video review (VR) and nonvideo review (NVR) groups on days 1, 15, and 30. †  $P < 0.02$  for both VR and NVR groups comparing days 1 and 15 by Wilcoxon test. ‡  $P < 0.01$  for VR group only comparing days 15 and 30 by Wilcoxon test.

NS = not significant.

had improved by at least three points in all skills evaluated. In contrast, residents in the NVR group improved by at least three points in only four skills (allowing the povidone iodine sufficient time to dry, inserting catheters, withdrawing needles, and securing catheters).

## Discussion

In the current study, we evaluated two separate measures of resident performance. First, each of the predetermined specific criteria were scored and, second, a general “overall” grade was assigned to each resident by each judge. These grades of 0–10 were based on the judges’ subjective evaluations of the residents general performance and were independent of the criteria scores. Fixed criteria were not applied to this subjective “overall” grade; nonetheless, there was excellent correlation (Pearson coefficients of 0.933, 0.846, and 0.799 on days 1, 15, and 30, respectively) between the individual summed scores based on specific objective criteria and the overall subjective grade. Furthermore, in many training programs, proficiency is determined by an overall general assessment rather than by specific fixed criteria.

**Table 5. Proportions\* (Percents) of Residents with at Least a 3-Point Improvement in Skills between Days 1 and 30 by Treatment Group**

|                        | VR        | NVR       |
|------------------------|-----------|-----------|
| Proper positioning     | 6/6 (100) | 5/6 (83)  |
| Aseptic technique      | 8/8 (100) | 3/7 (43)  |
| No spillage            | 8/8 (100) | 4/5 (80)  |
| Povidone iodine dried  | 8/8 (100) | 6/6 (100) |
| Draping                | 8/8 (100) | 6/8 (75)  |
| Needle control         | 8/8 (100) | 3/6 (50)  |
| Midline location       | 6/6 (100) | 2/3 (67)  |
| Cephalad direction     | 6/7 (86)  | 3/5 (60)  |
| Air amount             | 7/7 (100) | 5/6 (83)  |
| Catheter insertion     | 4/4 (100) | 6/6 (100) |
| Needle withdrawal      | 8/8 (100) | 8/9 (89)  |
| Catheter securing      | 8/8 (100) | 4/4 (100) |
| Patient treated gently | 6/7 (86)  | 4/7 (57)  |

\* The denominator of each skill is the number of residents whose scores could improve by at least 3 points.

VR = video review; NVR = nonvideo review.

In fact, in our study, although somewhat arbitrary, we left it up to the judge to decide what was a minor or major error in technique. Although having standardized measures for detecting a major *versus* minor error in each skill would have been optimal, it does not reflect current practice because no valid operational tools exist to distinguish between a major and minor error. Nonetheless, there was excellent interrater reliability (kappa coefficients) among the judges. The potential for bias, although unlikely, could not be excluded because, by study design, the attending using the videotape to teach and supervise the self-assessment was unblinded. Nonetheless, it is unlikely that residents in the NVR group were given inferior teaching because they had access to a variety of teaching resources and staff. The supervising attending physician and the judges were blinded to resident group.

Most residents improved their skills during the 30 days of instruction, and none performed progressively worse during the rotation. However, our data indicate that residents in the VR group achieved higher overall grades by the end of their rotation and improved to a greater degree than did residents in the NVR group. Although both groups of residents increased in skills between days 1 and 15, residents in the VR group appeared to continue to improve their skills as the month progressed. It has recently been suggested that manual dexterity, eye-hand coordination, and other motor abilities may be important determinants of an individual’s performance at obstetric epidural procedures, particularly after the initial training phase.<sup>8</sup> Acquiring skills in the performance of the epidural procedure requires psychomotor coordination, and an operator’s visual assessment of performance may play a key role in providing the feedback necessary to master a skill.<sup>9</sup> This might explain the apparent differences between the two groups during the second 2 weeks of the rotation, even though each resident had already completed more than 30 epidural procedures. Development of some skills, such as adherence to aseptic technique and needle control, appear to be facilitated with video review. The reason for this is unclear. Furthermore, the potential impact of self-assessment made possible through video review by a resident may have an important effect on learning.

Some residents in the current study began their first obstetric anesthesia rotation having previously acquired better epidural analgesia technique than others. Residents were randomly assigned to either the VR or the NVR group, and there was no significant difference in the rank of the grades between the two groups on day 1 (Mann–Whitney U test). Because our study evaluated change in skill from day 1 to day 30 separately for each group, any apparent differences on day 1 between groups are of minor impact. The denominator for calculating the percentage of residents with at least a three-point improvement of skills only included residents who

could improve their skills by that criterion. These were also evenly distributed between the VR and NVR groups.

As determined by the kappa coefficients in the current study, the interrater reliability among the attending physicians evaluating the videos was good to excellent. This could be attributable to the fact that the attendings involved in the study were seasoned anesthesiologists experienced in epidural analgesia technique. It is possible that interrater reliability may not be as good with a more heterogeneous group of attending physicians. It is also likely that residents in the VR group performed better because they had the opportunity to scrutinize their performance in greater detail and as often as necessary compared with residents in the NVR group. A "halo effect" resulting from a subject's awareness that he or she is being observed is probably negligible in the current study because both groups were videotaped and had similar scores on day 1. In addition, all residents knew that their films would be viewed whether or not they themselves were in attendance at the viewing, and all residents were supervised during placement of epidural analgesia by an attending anesthesiologist, as is the policy in our department.

Some authorities consider that epidural anesthesia technique can be rapidly mastered by residents and does not require elaborate teaching methods. Kopacz *et al.*<sup>10</sup> have suggested that significant improvement in technique occurs after the performance of just 25 epidural analgesia placement procedures. Although our study evaluated technique rather than success, the improvement observed in our residents after 2 weeks (and performance of approximately 30 epidural analgesia placement procedures) appears to support their observation. However, there are several factors particular to the obstetric setting that can make teaching of epidural analgesia placement more difficult than in the orthopedic or general surgical patient. The mother is awake, anxious, not medicated, and often in severe pain, which might cause her to move during the epidural analgesia placement. In addition, the woman's support person may be allowed in the labor room and may be watching the procedure. In many programs, a number of different attending physicians rotate through the labor and delivery ward on a daily basis, decreasing the continuity of resident training. Videotaping and replay do not require terribly expensive or sophisticated equipment, allow for much more careful analysis of technique, and by allowing visualization of a mistake in replay and slow motion, prevent the denial that often follows the suggestion by a

supervising attending during the procedure that a breach in technique has occurred.

After each review session, residents were asked how they felt about being videotaped. All stated that they felt somewhat uncomfortable with being taped on day 1, but none was uncomfortable by day 30. Residents who reviewed their videotapes suggested that videotaping motivated them to improve their technique.

A limitation of this teaching tool may be the cost associated with purchase of equipment and allocation of adequate faculty time for appropriate review of the tapes. Currently, the necessary equipment can be purchased for less than \$1,000. As with other new teaching modalities, such as simulators, video teaching as described in this study requires a significant commitment of faculty time.

In conclusion, under the conditions of the current study, videotaping and video review of residents initiating epidural analgesia on the labor and delivery ward resulted in greater improvement in overall and selected performance criteria than that of a group that did not have video review. Videotaping may help in acquiring epidural skills and may prove to be a valuable tool in training and motivating anesthesiology residents. Video review also permits teaching at a later time without heuristic and critical discussions in the presence of an awake parturient and support person. However, these conclusions may be affected by the fact that this was a relatively small study, and the findings may not necessarily apply to every residency program.

## References

1. Keats AS: Quality anesthesia care: A model of future practice of anesthesiology. *ANESTHESIOLOGY* 1977; 47:488-9
2. Miller G, Gabbard C: Effects of visual aids on acquisition of selected tennis skills. *Perception Motor Skills* 1988; 67:603-6
3. Cauraugh JH, Martin M, Martin KK: Modeling surgical expertise for motor skill acquisition. *Am J Surg* 1999; 177:331-6
4. McAvoy BR: Teaching clinical skills to medical students: The use of simulated patients and videotaping in general practice. *Med Educ* 1992; 22:193-9
5. Olsen JC, Gurr ED, Hughes M: Video analysis of emergency medicine residents performing rapid-sequence intubations. *J Emerg Med* 2000; 18:469-72
6. Sivarajan M, Lane PE, Miller EV, Liu P, Herr G, Willenkin R, Winter P, Hardy C, Mulroy MF: Performance evaluation: Continuous lumbar epidural anesthesia skill test. *Anesth Analg* 1981; 60:543-7
7. Fleiss JL: *Statistical Methods for Rates and Proportions*. New York, John Wiley & Sons, 1981, p 218
8. Dashfield AK, Coghill JC, Langton JA: Correlating obstetric epidural anaesthesia performance and psychomotor aptitude. *Anaesthesia* 2000;55:744-9
9. Rogers DA, Regeher G, Howdieshell TR, Yeh KA, Palm E: The impact of external feedback on computer-assisted learning for surgical technical skill training. *Am J Surg* 2000; 179:341-3
10. Kopacz DJ, Neal JM, Pollock JE: The regional anesthesia "learning curve": What is the minimum number of spinal blocks to reach consistency? *Reg Anesth* 1996; 21:182-90