Can J Anesth/J Can Anesth (2010) 57:120–126
DOI 10.1007/s12630-009-9219-2

REPORTS OF ORIGINAL INVESTIGATIONS

Anesthesiologists' learning curves for ultrasound assessment of the lumbar spine

Les courbes d'apprentissage des anesthésiologistes en matière d'évaluation de la colonne lombaire par échoguidage

Clarita B. Margarido, MD, PhD · Cristian Arzola, MD · Mrinalini Balki, MD · Jose C. A. Carvalho, MD, PhD

Received: 10 September 2009 / Accepted: 30 October 2009 / Published online: 30 December 2009

Abstract

Background and objectives *Ultrasound assessment of the lumbar spine to facilitate neuraxial anesthesia has recently received much attention. The transfer of knowledge pertaining to this skill has never been studied. The purpose of this study was to determine the amount of teaching needed to achieve competence in spinal ultrasound.*

Methods *Participants were given reading material and a link to a video presentation on spinal ultrasound. As well, they attended a 45-min lecture followed by a 30-min hands-on workshop. They were then assessed individually 1–2 weeks later. The assessment was performed on a live model using a low frequency curved ultrasound probe (2–5 mHz). Participants were asked to determine, at random lumbar spinal levels, the optimal insertion point and the depth to the ligamentum flavum-dura mater unit (up to 20 trials, 2 min per trial). Feedback was provided by an expert for each incorrect task. The learning curves were constructed, using the cumulative sum method, by comparing the participants' results with those of a benchmark established by experts. Statistical analysis was performed using STATA® 9.2 for Macintosh (College Station, TX, USA).*

Results *A total of 308 anesthesiologists were approached and 18 anesthesiologists participated in the study. Only five of the 18 participants (27%) achieved competence in determining the interspace, with a median number of 11 attempts (range 8–18). None of the participants achieved competence in determining either the insertion point or the depth to the ligamentum flavum-dura mater unit.*

Conclusions *Under the study conditions, 20 supervised trials plus teaching sessions were not enough for the participants to achieve competence in different aspects of ultrasound assessment of the lumbar spine. These results may well be considered when planning teaching sessions and workshops in the future.*

Résumé

Contexte et objectifs *L'évaluation de la colonne lombaire par échoguidage pour faciliter l'anesthésie neuraxiale a récemment fait l'objet d'une attention soutenue. Le transfert des connaissances liées à cette aptitude n'a jamais été étudié. L'objectif de cette étude était de déterminer la quantité d'enseignement requise pour acquérir les compétences nécessaires à l'échoguidage rachidien.*

Méthode *Les participants à l'étude ont reçu du matériel à lire ainsi qu'un lien pour accéder à une présentation vidéo portant sur l'échoguidage rachidien. De plus, ils ont assisté à un cours de 45 minutes suivi par un atelier pratique de 30 minutes. Puis ils ont été évalués individuellement 1-2 semaines plus tard. L'évaluation a été réalisée sur un modèle vivant à l'aide d'une sonde d'échoguidage courbe à basse fréquence (2-5 mHz). On a demandé aux participants d'identifier, à différents niveaux de la colonne lombaire choisis aléatoirement, le point d'insertion optimal et la profondeur jusqu'à l'unité ligament jaune – dure-mère (jusqu'à 20 essais, deux minutes par essai). Un expert commentait chacune des tâches incorrectes. Les courbes d'apprentissage ont été*

C. B. Margarido, MD, PhD · C. Arzola, MD · M. Balki, MD
Department of Anesthesia and Pain Management, Mount Sinai Hospital, University of Toronto, Toronto, ON, Canada

J. C. A. Carvalho, MD, PhD (✉)
Department of Anesthesia and Pain Management, Mount Sinai Hospital, University of Toronto, 600 University Avenue, Room 781, Toronto, ON M5G 1X5, Canada
e-mail: jose.carvalho@uhn.on.ca

réalisées à l'aide de la méthode CUSUM en comparant les résultats des participants à ceux d'un jalon établi par des experts. L'analyse statistique a été réalisée à l'aide de STATA® 9.2 pour Macintosh (College Station, TX, USA).
Résultats *Au total, 308 anesthésiologistes ont été approchés et 18 ont pris part à l'étude. Seuls cinq des 18 participants (27 %) ont acquis la compétence nécessaire à identifier l'espace intervertébral, avec un nombre moyen de 11 tentatives (fourchette 8-18). Aucun des participants n'a démontré la compétence nécessaire à déterminer le point d'insertion ou la profondeur de l'unité ligament jaune – dure-mère.*
Conclusion *Dans les conditions de l'étude, 20 essais supervisés et des sessions d'enseignement n'ont pas suffi pour que les participants acquièrent les compétences nécessaires à réaliser différents aspects de l'évaluation de la colonne lombaire par échoguidage. Ces résultats devraient à l'avenir être pris en compte lors de la planification de sessions d'enseignement et d'ateliers.*

Introduction

Ultrasound imaging has become an increasingly popular skill among anesthesiologists to facilitate vascular access, peripheral nerve blocks, and neuraxial anesthesia.[1–6]

The success of neuraxial anesthetic techniques depends ultimately on a very accurate assessment of the spine. Currently, the most common method of assessment is palpation of anatomic landmarks, particularly the iliac crests and the spinous processes. Unfortunately, even when these landmarks are palpable, this method is highly inaccurate, particularly for the assessment of the level of puncture.[7] Spinal ultrasound has proven to be a valuable clinical tool to improve the accuracy of such assessments as well as an important teaching tool in neuraxial anesthesia.[8]

However, in most studies related to ultrasound-facilitated neuraxial anesthesia, the scanning was performed by the few anesthesiologists regularly practicing ultrasound imaging. Whether this technique can be learned and implemented by novice anesthesiologists as well as other specialists is yet to be determined.

There are only a few reports in the literature addressing the teaching of ultrasound-facilitated regional techniques, and most pertain to needling techniques and peripheral nerve blocks.[8–10] Grau *et al.* showed that the learning curve for labour epidural placement improved significantly when ultrasound imaging was used for teaching epidural anesthesia in obstetrics.[8] In their study, however, the trainees were offered ultrasound information on the patients before placing the epidurals, but did not themselves perform the ultrasound scanning. Therefore, the degree of difficulty of transferring the requisite knowledge to perform spinal ultrasound remains unclear.

The purpose of this study was to determine the amount of training, especially hands-on experience, needed for an anesthesiologist to achieve the requisite competence in spinal ultrasound assessment to perform ultrasound-facilitated spinal and epidural anesthesia.

Methods

The study was designed as a prospective cohort investigation, and was conducted from September to December 2008. Following approval by the Research Ethics Board of Mount Sinai Hospital, 308 anesthesiologists, whose names were included in a database used to promote our departmental annual meetings, were approached by e-mail. The eligible participants were anesthesiologists with no previous spinal ultrasound experience who were capable of performing spinals and epidurals independently. The participants who enrolled in the study gave their written informed consent.

Teaching intervention

Participants were provided reading material and an educational video on ultrasound-facilitated spinals and epidurals,[1] which they were required to review prior to proceeding to the next teaching step. One week later, they attended a 45-min lecture on the topic, followed by a 15-min demonstration on a live model and a 30-min hands-on workshop in a small group setting. According to the expert investigators, the live models chosen had easily palpable spines, no visible spinal abnormalities, and typical sonoanatomy at all the lumbar spaces. During the workshop, the participants received feedback on their performances and were instructed and prepared to perform the following ultrasound scanning tasks independently: (1) in the longitudinal paramedian plane, identification of the sacrum and five lumbar intervertebral spaces (IS); (2) in the transverse plane, determination of the interspace and the midline (consequently the optimal insertion point [IP]) at different lumbar interspaces; and (3) determination of the distance from the skin to the inner aspect of the ligamentum flavum-dura mater unit (LF).

With the models in the sitting position, ultrasound scanning was performed using a portable MicroMaxx® ultrasound system equipped with a 2–5 mHz curved array probe (Sonosite® Canada Inc., Markham, ON, Canada). The best possible image of the sacrum was captured by positioning the probe along the long axis of the lumbar spine for

---

[1] *Carvalho JCA*. Ultrasound-guided epidural anesthesia video tutorial. Available from URL: http://sonositelearning.com/shop/ultrasound_guided_epidural_ane.php.

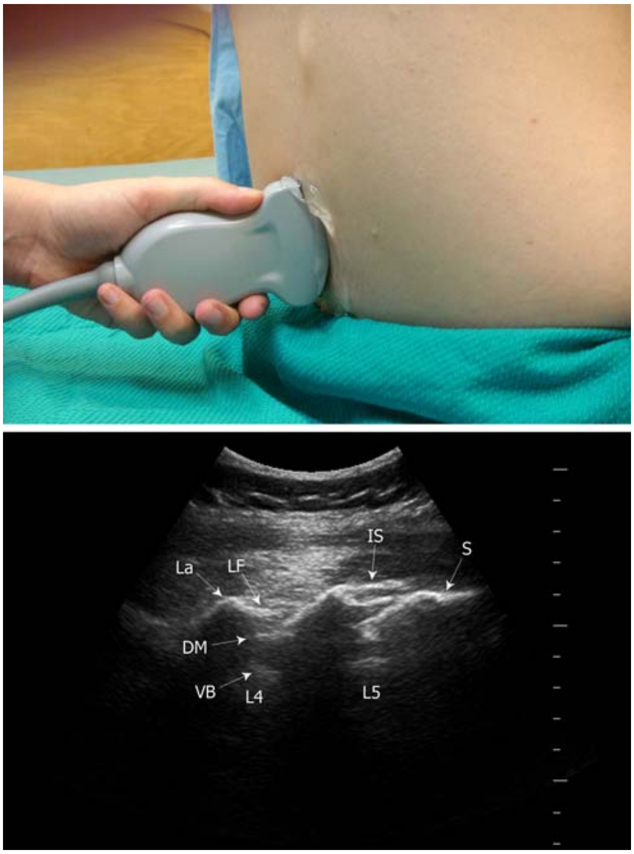

Fig. 1 *Top*: Curvilinear probe is positioned in the longitudinal paramedian plane over the sacrum and the lower lumbar spine. *Bottom*: Typical sonogram depicting the sacrum (S), the laminas (La) of L5 and L4, the vertebral bodies (VB) of L5 and L4, and the L5–S1 and L4–L5 interspaces (IS), with the ligamentum flavum (LF) and posterior dura mater (DM)

the longitudinal paramedian approach (Fig. 1). While in the longitudinal plane, the probe was then moved cephalad to determine the desired IS. When the desired IS was ascertained, the probe was positioned in the transverse plane over the identified IS to determine the optimal IP and the depth to the LF. The midline of the spine corresponding to the spinous process was identified as a hyperechoic signal underneath the skin, continuing as a hypoechoic vertical shadow (Fig. 2). A mark representing the midline was drawn on the skin at the midpoint of the upper surface of the probe. The probe was then moved slightly cephalad or caudad to allow a detailed view of the IS, which included the transverse and articular processes, the LF, and the vertebral body (Fig. 2). When the best image of the IS was captured in the transverse plane, a second mark was drawn on the skin at the midpoint of the lateral aspect of the probe. The optimum IP was determined by intersecting an imaginary vertical line through the first skin mark with an imaginary horizontal line through the second skin mark. The image of the IS was frozen and the distance from the skin to the LF was measured with the aid of a built-in calliper.

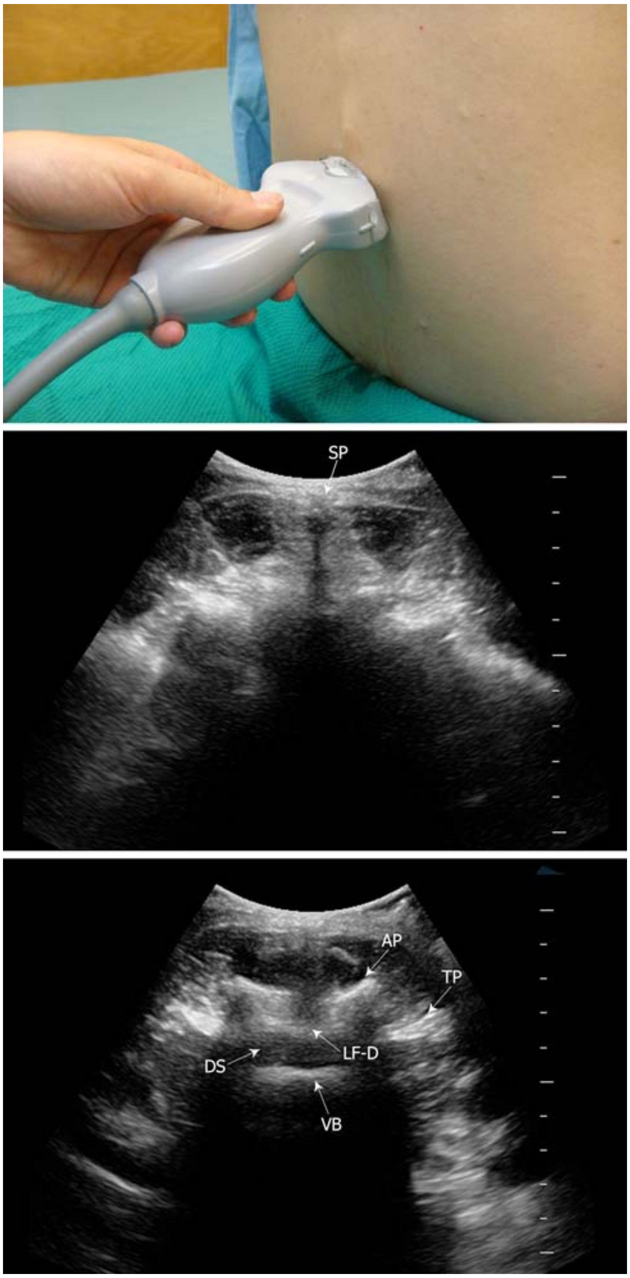

Fig. 2 *Top*: Curvilinear probe is positioned in the transverse plane, either on the tip of the spinous process or on the interspace. *Middle*: Probe positioned on the tip of the spinous process (SP) generates a hyperechoic image and a vertical hypoechoic acoustic shadow. *Bottom*: Probe positioned on the interspace depicts the articular processes (AP), the transverse processes (TP), the ligamentum flavum-posterior dura mater unit (LF-D), the vertebral body (VB) and the dural sac (DS)

Assessment

Seven to 14 days following the teaching session, the participants were assessed individually on their scanning skills and their ability to perform specific tasks. We did not control for the participants using the ultrasound in their practice during this time. The assessment required participants to perform the same tasks with the same equipment as during the teaching sessions and to use a live model with an easy spine.

Prior to the participants beginning their assessments, each model's lumbar area was marked with 30 cm lines on the *x* and *y* axes, and a transparent template was applied to the skin over the lines (Fig. 3). Three expert anesthesiologists with more than three years of experience in spinal ultrasound (C.A., M.B., and J.C.) scanned the models' lumbar spines to determine the optimal IP (coordinates marked on the *x* and *y* axes) and depth to the LF for all five lumbar interspaces. The experts were blinded to each others' assessments. The average of the three assessments was considered as a benchmark and subsequently used to assess the performance of the participants.

The participants were asked to determine the IS, the IP, and the distance from the skin to the LF at random lumbar interspaces. Each participant was asked to perform these three tasks up to 20 times, with a maximum of 2 min allowed per attempt, and a maximum of 50 min allowed per participant, including necessary performance feedback. For each trial, a different IS was selected, and the sequence of the IS was randomly determined. Feedback was provided by an expert for each incorrect task. The midline and the midpoint of the interspace at each of the interspaces (determinants of the IP) were marked on the transparent template on the *x* and *y* axes (Fig. 3). A new template was applied to the models' lumbar areas for each sequence of five IS.

Fig. 3 Assessment of the insertion point: transparent templates are positioned on the back of the model over the *x* and *y*-axes. For each task, the midline and the interspace are marked on the template by the experts and also by the participants according to precise instructions

Outcomes

The primary outcomes of the study were the number of attempts required by the participants to correctly identify the IS, to determine the optimal IP, and to measure the distance from the skin to the LF. The identification of the IS was assessed as a binary outcome being classified as correct or incorrect. The IP and the distance to the LF were compared with the benchmarks of the experts. The difference (in mm) between the assessment of these parameters by the participant and the average of the three experts was noted. The IP and the depth to the LF were considered correct if they were within 5 mm of the benchmark. These limit values were defined based on the maximum deviation found among the experts' assessments.

Statistical analysis and sample size calculation

In order to construct learning curves for the primary outcomes, the series of attempts performed by the participants were analyzed by the cumulative sum (Cusum) graphical method.[11] On a Cusum graph, the cumulative differences of the quality characteristic from a target level are plotted in sequence, allowing detection of deviations from predetermined standards. Variables for the construction of a Cusum graph are the acceptable (p_0) and the unacceptable (p_1) failure rates and reasonable probabilities of type I and II errors (α and β). From these, two decision limits (h_1 and h_0) and the variable s are calculated (Appendix 1). The graph starts at zero (Cusum score). For each successful attempt, the amount s is subtracted from the previous Cusum score. For each failed attempt, the amount 1-s is added to the previous Cusum score. Thus, a negative trend of the Cusum line indicates success, whereas a positive trend indicates a failure at the procedure being analyzed. When the line crosses the upper decision limit (h_1) from below, the actual failure rate is significantly greater than the acceptable failure rate, with a probability of a type I error equal to α. In terms of performance, the subject is deemed not to have achieved competence in the series of attempts. When the line crosses the lower decision limit (h_0) from above, the true failure rate does not differ significantly from the acceptable failure rate, with a probability of a type II error equal to β. In this case, the subject is deemed to have achieved competence in the series of attempts. When the Cusum line is between the decision limits, no statistical inference can be made and the performance remains undetermined, indicating that more attempts are necessary.

The null hypothesis indicates that the actual failure rate is not different from the acceptable failure rate; on the other hand, the alternative hypothesis expresses the actual failure rate as equal to or exceeding the unacceptable failure rate. For the sample size calculation of our study, α and β were set to 10%; whereas acceptable and unacceptable failure rates were set to 20 and 40%, respectively, based on the opinion of the aforementioned three spinal ultrasound experts. With these variables, it was expected that participants would require from 17 to 19 attempts to achieve competence at the proposed tasks. Statistical analysis was performed with Stata® 9.2 for Macintosh (College Station, TX, USA).

Results

A total of 18 of the 308 anesthesiologists who were approached participated in the study. In the pre-determined period of 50 min per participant, 11 participants were able to accomplish 20 attempts; four participants were able to accomplish 15 attempts; one participant was able to accomplish 12 attempts; and two participants were only able to accomplish 10 attempts.

Interspace

In the Cusum learning curves for the IS task (Fig. 4), it was demonstrated that only five of 18 participants (27%) achieved competence with a median number of 11 attempts (range from 8 to 18). Four participants did not achieve competence, and the performance remained undetermined in nine participants, suggesting the need for more attempts to assess competence.

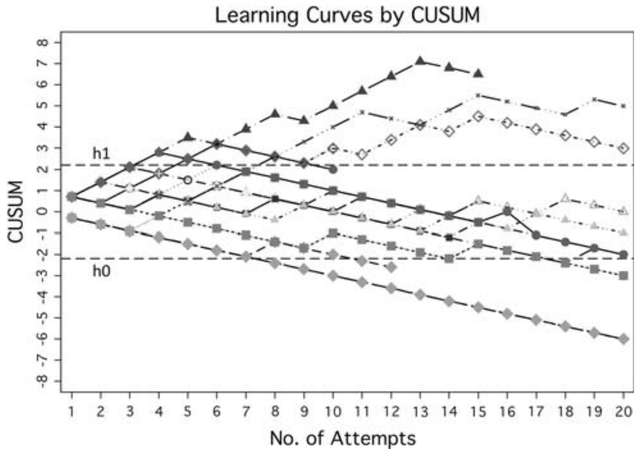

Fig. 4 Learning curves for Interspace by Cusum graph. Lines represent learning curves for individual participants. Upper and lower decision limits are represented by dashed lines h_1 (2.2) and h_0 (−2.2), respectively. Only five participants crossed h_0 from above to achieve competence

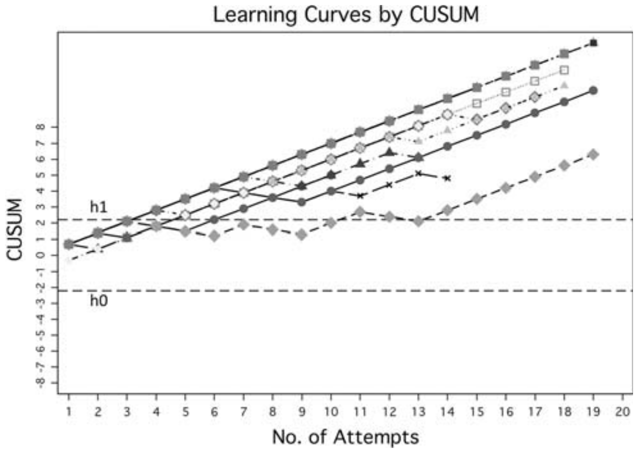

Fig. 5 Learning curves for Insertion Point by Cusum graph. Lines represent learning curves for individual participants. Upper and lower decision limits are represented by dashed lines h_1 (2.2) and h_0 (−2.2), respectively. All the participants crossed the line h_1 from below, indicating that the actual failure rates are significantly greater than the acceptable failure rate (20%), and in terms of performance, they are deemed not to have achieved competence in the series of attempts

Insertion point

None of the participants achieved competence at determining the optimal IP (Fig. 5). The overall median success rate was 6.35% (percentile 25-percentile 75, 0–15.8), which is much lower than the acceptable failure rates that we set for the Cusum analysis.

Distance from the skin to the LF

In the analysis of learning curves by Cusum (Fig. 6), four participants did not achieve competence and 14 participants had undetermined performances.

Discussion

This study was conducted to explore the efficacy and difficulties of transferring the knowledge of ultrasound assessment of the lumbar spine in order to facilitate the placement of epidural and spinal anesthesia.

For anesthesiologists, acquiring the various skills related to using ultrasound for vascular access and regional anesthesia may depend on the complexity of the skill, and may require varying degrees of teaching or training to achieve competence. De Oliveira Filho *et al.* determined learning curves while inserting needles into a bovine muscular

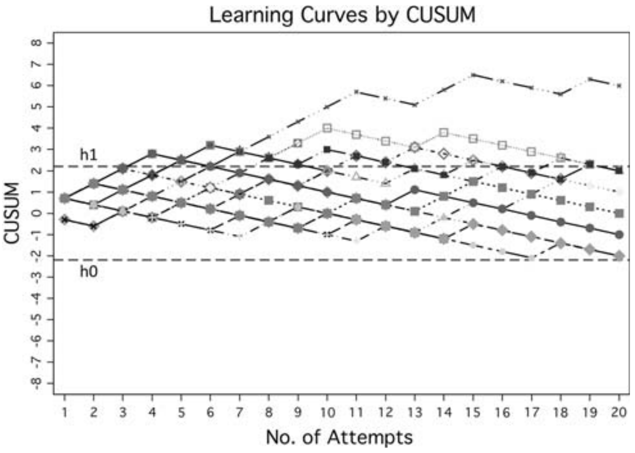

Fig. 6 Learning curves for Depth to the LF by Cusum graph. Lines represent learning curves for individual participants. Upper and lower decision limits are represented by dashed lines h_1 (2.2) and h_0 (−2.2), respectively. None of the participants crossed h_0 from above to determine competence in the series of attempts. Four participants crossing h_1 from below did not achieve competence and, in the majority of the participants (14/18), the lines were between the decision limits, consequently not providing a statistical inference concerning competence and indicating that more attempts were required to assess the performance

phantom in plane with a linear transducer and also while attempting to use the same skill to hit a certain target.[11] They estimated the number of attempts to achieve a 95% success rate in these two skills to be 39 and 109, respectively.

We hypothesized that spinal ultrasound assessment, not being a real-time procedure and requiring less hand-eye coordination, would be an easier skill to teach and learn. Apparently, this was not the case. We found that determining the interspace was the only task that could be accomplished with consistency by some participants, unlike determining the optimal IP and the depth to the LF. Although the recognition of ultrasound patterns is essential for determining the IP and the distance to the LF, it is certainly not enough. Insertion point accuracy relies on making precise skin markings while holding the probe steadily; inattention while marking the exact midpoints of the probe and probe movements could explain the results found in our study. Similarly, the measurement of the distance from the skin to the LF requires attention to detail, including timely freezing of the image and precise use of the built-in calliper. It was our impression during the study that the participants recognized the structures but were not meticulous in completing their tasks. In the case of identifying the interspace, we found that the L5–S1 interspace was sometimes missed by the participants, and that led to incorrect performance of the task and an impaired learning curve. In the case of determining the IP, two problems were identified: (1) While moving the probe from the longitudinal plane to the transverse plane, participants sometimes slid it to a different interspace, especially in the lower lumbar interspaces where the distance between two IS is smaller; (2) When the space was correct, usually participants were not meticulous in marking the interspace as instructed, i.e., where the skin mark should coincide with the exact position of the 1-mm wide ultrasound beam generated from the probe. Finally, in the case of the distance to the LF, the same explanations apply, i.e., either the space was no longer the one assigned or the calliper was not meticulously positioned.

It is well established that videos and live demonstrations are both important to acquiring proficiency in manual skills.[12–14] In our study, we provided reading material, video instruction, a lecture, and a hands-on workshop in which each participant was assisted in performing the required tasks correctly. The teaching experience was more comprehensive than those usually offered in workshops at various meetings. Although our group has conducted many ultrasound workshops, we have noted the difficulty in assessing the performance of the participants and, most importantly, their ability to retain the knowledge. The results of our study show that a more detailed teaching plan and instruction additional to that currently offered are necessary if the goal is to transfer knowledge effectively. Perhaps our results will contribute to the re-design of our approach to conducting workshops.

Our study has some limitations. First, even though the workshops were conducted with live models, they were models and not patients. We acknowledge that the performance of spinal ultrasound on live models differs from a real-life situation; however, logistical limitations and the vast number of attempts required for the study prevented us from using patients. Second, the study only evaluates the teaching of spinal assessment with ultrasound and not how that information impacts the performance of actual neuraxial procedures.

In conclusion, in addition to the educational material, achieving competence in spinal ultrasound assessment requires a lecture, a demonstration workshop, and more than 20 supervised trials. The determination of the IP and the distance from the skin to the LF seem to be the most difficult skills to acquire. These findings should be useful in planning future workshops, as the current ones are generally restricted to limited hands-on experience. In addition, while teaching hands-on workshops, special attention should be given to the meticulous marking and measurement of landmarks and distances.

Acknowledgement The authors thank Kristi Downey, Perinatal Anesthesia Research Coordinator, Department of Anesthesia and Pain Management, Mount Sinai Hospital, 600 University Avenue, Room 781, Toronto, Ontario M5G 1X5, Canada.

Funding Departmental.

Conflicts of interest None of the authors has financial or personal relationships or affiliations that could influence (or bias) this research.

Appendix 1

Calculation of sample size (number of attempts):

1. Average number of attempts with acceptable failure rate:

   $[(h_0(1-\alpha) - \alpha h_1)/(s-p_0)]$

2. Average number of attempts with unacceptable failure rate:

   $[(h_1(1-\beta) - \beta h_0)/(p_1-s)]$

$\mathrm{a} = \ln[(1-\beta)/\alpha]$

$\mathrm{b} = \ln[(1-\alpha)/\beta]$
$\mathrm{P} = \ln\,(p_1/p_0)$
$\mathrm{Q} = \ln[(1-p_0)/(1-p_1)]$
$S = \mathrm{Q}/(\mathrm{P}+\mathrm{Q})$
$h_0 = -\mathrm{b}/(\mathrm{P}+\mathrm{Q})$
$h_1 = \mathrm{a}/(\mathrm{P}+\mathrm{Q})$

References

1. *Grau T*, *Leipold RW*, *Conradi R*, *Martin E*, *Motsch J*. Efficacy of ultrasound imaging in obstetric epidural anesthesia. J Clin Anesth 2002; 14: 169–75.
2. *Augoustides JG*, *Horak J*, *Ochroch AE*, *et al.* A randomized controlled clinical trial of real-time needle-guided ultrasound for internal jugular venous cannulation in a large university anesthesia department. J Cardiothorac Vasc Anesth 2005; 19: 310–5.
3. *Chantzi C*, *Saranteas T*, *Paraskeuopoulos T*, *Dimitriou V*. Ultrasound and transcutaneous neurostimulator combined technique as a training method for nerve identification in anesthesia residents. Reg Anesth Pain Med 2007; 32: 365–6.
4. *Arzola C*, *Davies S*, *Rofaeel A*, *Carvalho JC*. Ultrasound using the transverse approach to the lumbar spine provides reliable landmarks for labor epidurals. Anesth Analg 2007; 104: 1188–92.
5. *Popovic J*, *Morimoto M*, *Wambold D*, *Blanck TJ*, *Rosenberg AD*. Current practice of ultrasound-assisted regional anesthesia. Pain Pract 2006; 6: 127–34.
6. *Carvalho JC*. Ultrasound-facilitated epidurals and spinals in obstetrics. Anesthesiol Clin 2008; 26: 145–58.
7. *Broadbent CR*, *Maxwell WB*, *Ferrie R*, *Wilson DJ*, *Gawne-Cain M*, *Russell R*. Ability of anaesthetists to identify a marked lumbar interspace. Anaesthesia 2000; 55: 1122–6.
8. *Grau T*, *Bartusseck E*, *Conradi R*, *Martin E*, *Motsch J*. Ultrasound imaging improves learning curves in obstetric epidural anesthesia: a preliminary study. Can J Anesth 2003; 50: 1047–50.
9. *Sites BD*, *Gallagher JD*, *Cravero J*, *Lundberg J*, *Blike G*. The learning curve associated with a simulated ultrasound-guided interventional task by inexperienced anesthesia residents. Reg Anesth Pain Med 2004; 29: 544–8.
10. *Sites BD*, *Spence BC*, *Gallagher J*, *et al.* Regional anesthesia meets ultrasound: a specialty in transition. Acta Anaesthesiol Scand 2008; 52: 456–66.
11. *de Oliveira Filho GR*, *Helayel PE*, *da Conceição DB*, *Garzel IS*, *Pavei P*, *Ceccon MS*. Learning curves and mathematical models for interventional ultrasound basic skills. Anesth Analg 2008; 106: 568–73.
12. *Wong DT*, *Prabhu AJ*, *Coloma M*, *Imasogie N*, *Chung FF*. What is the minimum training required for successful cricothyroidotomy? A study in mannequins. Anesthesiology 2003; 98: 349–53.
13. *Birnbach DJ*, *Santos AC*, *Bourlier RA*. The effectiveness of video technology as an adjunct to teach and evaluate epidural anesthesia performance skills. Anesthesiology 2002; 96: 5–9.
14. *Gaba DM*. Two examples of how to evaluate the impact of new approaches to teaching. Anesthesiology 2002; 96: 1–2.