Face and construct validity of TU-Delft epidural simulator and the value of real-time visualization

Zivkovic, Nenad; Van Samkar, Ganapathy; Hermanns, Henning; Lirk, Philipp; Hollmann, Markus W.; Van Den Dobbelsteen, John J.; Van Gerwen, Dennis J.; Stevens, Markus J.

DOI
10.1136/rapm-2018-100161

Publication date
2019

Document Version
Final published version

Published in
Regional Anesthesia and Pain Medicine

Citation (APA)
Zivkovic, N., Van Samkar, G., Hermanns, H., Lirk, P., Hollmann, M. W., Van Den Dobbelsteen, J. J., Van Gerwen, D. J., & Stevens, M. J. (2019). Face and construct validity of TU-Delft epidural simulator and the value of real-time visualization. *Regional Anesthesia and Pain Medicine*, *44*(3), 298-302. https://doi.org/10.1136/rapm-2018-100161

Face and construct validity of TU-Delft epidural simulator and the value of real-time visualization

Nenad Zivkovic,[1] Ganapathy van Samkar,[1] Henning Hermanns,[1] Philipp Lirk,[1] Markus W Hollmann,[1] John J van den Dobbelsteen,[2] Dennis J van Gerwen,[2] Markus F Stevens[1]

[1]Department of Anesthesiology, Amsterdam UMC, University of Amsterdam, Amsterdam, The Netherlands
[2]Technical Department of BioMechanical Engineering, Technical University of Delft, Delft, The Netherlands

Correspondence to
Dr Henning Hermanns, Department of Anesthesiology, Amsterdam UMC, University of Amsterdam, Anesthesiology, Amsterdam 1105, The Netherlands;
h.hermanns@amc.nl

This work was presented in part at Euroanaesthesia 2016, May 28–30, 2016, in London, UK.

Received 11 October 2018
Accepted 22 October 2018
Published Online First 22 January 2019

To cite: Zivkovic N, van Samkar G, Hermanns H, *et al*. *Reg Anesth Pain Med* 2019;**44**:298–302.

ABSTRACT

Background and objectives Learning epidural anesthesia traditionally involves bedside teaching. Visualization aids or a simulator can help in acquiring motor skills, increasing patient safety and steepening the learning curve. We evaluated the face and construct validity of the TU-Delft Epidural Simulator and the effect of needle visualization.

Methods Sixty-eight anesthesiologists, anesthesia residents, and final-year medical students tested the epidural simulator. Participants performed six epidural simulations with and six without needle visualization. We tested face validity on a Likert scale questionnaire. We collected data with the simulator software (spinal taps, dura contacts, bone contacts, attempts, and time) and tested for correlation with the performer's experience (construct validity). A visualization aid was tested in a randomized crossover design.

Results Face validity as rated by the participants was above average, with a mean of 3.7 (2.0–4.8) on a 5-point scale. Construct validity was indicated by significantly more spinal taps (0.4 [0–4) vs 0.07 [0–2], p=0.04) and more dura contacts (0.58 [0–6] vs 0.37 [0–3], p=0.002) by the inexperienced group compared with the expert group. The visualization aid improved performance by reducing the number of bone contacts and the number of attempts, and by decreasing the procedure time. Prior visualization training reduced the total procedure time from 279 s (69–574) to 180 s (53–605) (p=0.01) for the "blind" procedure.

Conclusions The TU-Delft Epidural Simulator is a useful tool for teaching motor skills during epidural needle placement. Prior use of a visualization tool improves performance even without visual support during consequent simulations.

INTRODUCTION

Epidural catheter placement requires motor skills and experience. These are generally acquired during hands-on training, subjecting the patient to (unnecessary) risks. Although a variety of teaching methods for gathering technical skills have been described, there is no widely accepted method to test the manual skills of anesthesiologists during epidural needle placement.[1] Simulators can provide a safe environment for teaching residents and can also be used as a valuable tool for assessing the resident's proficiency in a systematic and consistent way, before performing this procedure on patients. An extensive technical review of 31 different epidural and spinal simulators was done by Vaughan and colleagues,[2] comparing their features and highlighting their advantages and shortcomings. However, neither construct nor face validity was available in this review for any of the described simulators. More recently, a banana was suggested as a teaching model for loss of resistance (LOR) after comparison with three simulators.[3] Another study describing a recently developed simulator uses pressure guidance for detection of LOR but lacks the advantage of MRI modeling.[4] The TU-Delft Simulator for Epidural Needle Skills (SENS) with 2 degrees of freedom was used in our study. It has the advantage of modeling a variety of MRI scans. Thus, varied constitutions, anatomies, and possible pathologies of the vertebral column can easily be implemented in the simulator software. High-fidelity simulators have not proven to be superior to low-fidelity ones in terms of clinical impact. A study by Friedman and colleagues suggests no difference in the learning curve between residents taught on low-fidelity "greengrocer's model" and the ones taught on a high-fidelity simulator.[5 6] However, the study did not include a control group. High-fidelity simulators offer up to 6 degrees of freedom, allowing the user to choose insertion point and needle trajectory in all plains and axes, while low-fidelity simulators mimic only relevant clinical features. Which features are mandatory in a simulator and which are superfluous still remain a topic of discussion. The purpose of our study was to test the TU-Delft SENS considering three different issues: face validity (the relevance of a test as it appears to test participants), construct validity (the degree to which a test measures what it claims to be measuring), and the effect of a visualization aid. We hypothesized that participants would rate the simulator as realistic and useful for training purposes (mean score on the Likert scale >2.5). Furthermore, our test for construct validity was that experienced anesthesiologists would perform better than inexperienced residents or students. We expected decreased bone contacts, dura contacts, and spinal taps; we also expected fewer attempts and less time spent until completion of procedure. We regarded the number of spinal taps as being clinically most relevant. The other measures for safety and quality of performance were added because we expected the incidence of spinal taps to be too low to reach significance. Finally, we hypothesized that a visualization aid improves the performance during the actual procedure and in subsequent procedures without visualization.

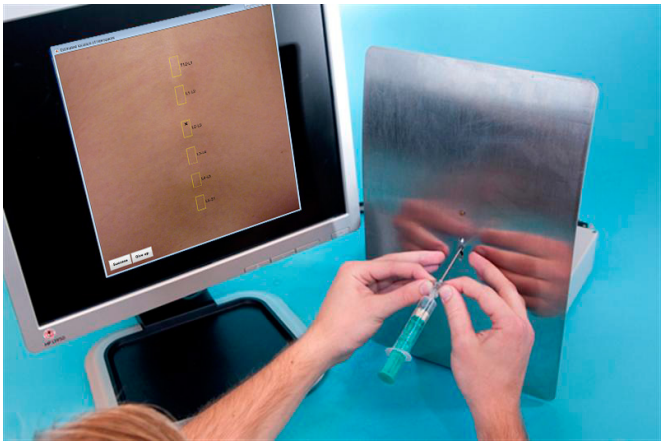

Figure 1 Hardware of the simulator showing the monitor, the metal plate with needle, and the syringe.

METHODS

The TU-Delft SENS is a computer-controlled learning tool that provides force feedback based on a virtual patient model. It consists of a metal plate representing the dorsal side of the patient, with a vertical opening. A syringe with needle can be introduced through this opening at a fixed point of entry. The needle can be rotated and can be angled at different angles in respect to the vertical plane. The needle can be advanced anterioposteriorly toward the virtual front of the model (patient). Once the needle is advanced, the angle cannot be changed. Data are acquired electronically through a computer. The LOR is felt through force feedback, and so is the bone contact. MRI simulation shows actual (real-time overlay) advancement of the needle in the MRI model of the spine. The simulator (figures 1 and 2) allows the trainee to insert a needle into a virtual patient's back, following the midline approach with the LOR technique. The simulator software offers a flexible insertion point on the computer screen with visible spinous processes, while the hardware has a fixed insertion point and offers 2 degrees of freedom. The needle can be angled with respect to the back (vertical plane), mimicking the insertion (first degree of freedom). The needle can also be advanced inward toward the epidural space (second degree of freedom). The simulator allows training of several features of epidural needle placement: selection of the needle insertion point and angle (in the sagittal plane), and insertion of the needle with variable resistance simulating fat, supraspinous ligament, interspinous ligament, bone, epidural space, dura, and the intrathecal space. Additionally it provides tactile identification of epidural space entry by LOR with air or saline. For didactic reasons, changing the angle of the needle after passing the supraspinous ligament is not possible. The forces model and the loss-of-resistance-pressure model are based on a combination of actual force and pressure measurements (porcine specimens, in vivo and in vitro), data from the literature, and expert opinion.[7–10] There are no experimental studies on the force measurement during real epidural needle insertion on live humans, and therefore a realistic force range has been determined from animal studies.[9] The technical set-up and exact mechanism of action of the TU-Delft simulator are described in detail in a previous study.[10]

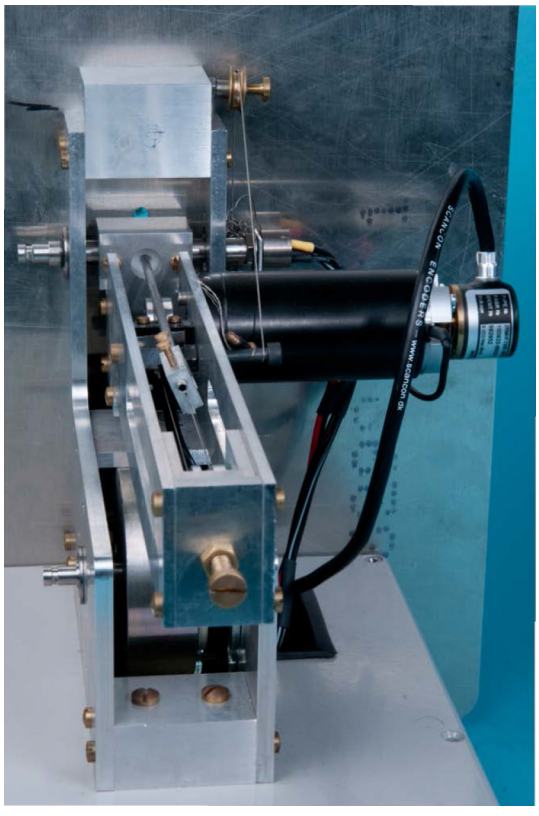

Figure 2 Details of the mechanical construction of the simulator, back side of the metal plate.

The anatomic model is based on segmented CT and MRI data. Although the system database contains anatomic models of 52 different patients, in this study we used a single patient model of six consecutive vertebral interspaces from T12–L1 to L5–S1 in order to keep these variables constant. The simulator software allows visual support to be displayed in the form of an MRI with representation of the needle during the procedure (figure 3). This allows the user to correct the angle of the needle if necessary, before contacting the bone. The optimal point and angle of insertion was recently studied in a computerized model.[11] This interface can be turned off or on.

Anesthesiologists, anesthesiology residents, and students of our department were included. Their experience in epidural needle placement varied from zero procedure to more than 1000. We divided the participants into two groups based on their experience. Novices were defined as having performed up to 30 epidural punctures, as suggested by the literature.[12–15] The experienced group was defined as those having performed more than 30 epidurals. Participants were asked to position/align the needle in the sagittal plane and then to insert it into the epidural space along a straight line. On reaching the epidural space, the participants gave oral confirmation and proceeded with the next interspace until completion of the study task. Thus, each participant performed 12 epidural needle placements in total, 6 with and 6 without the visual support turned on. Whether the participant started with the visual support turned on or off was decided by computer randomization. We randomized participants to either performing the epidurals *with simultaneous needle visualization* on MRI or to *first perform the punctures blindly*. In this manner, the value of a prescan visualization aid was evaluated. All participants received a standardized introduction to the simulator that included the content and features of the simulator and an explanation of the study questions. Participants were informed that the purpose of this study was to evaluate the simulator and not the participants, and that all data were saved anonymously. After performing 12 epidural needle placements, participants completed a form consisting of 11 questions to be answered on a modified Likert scale.

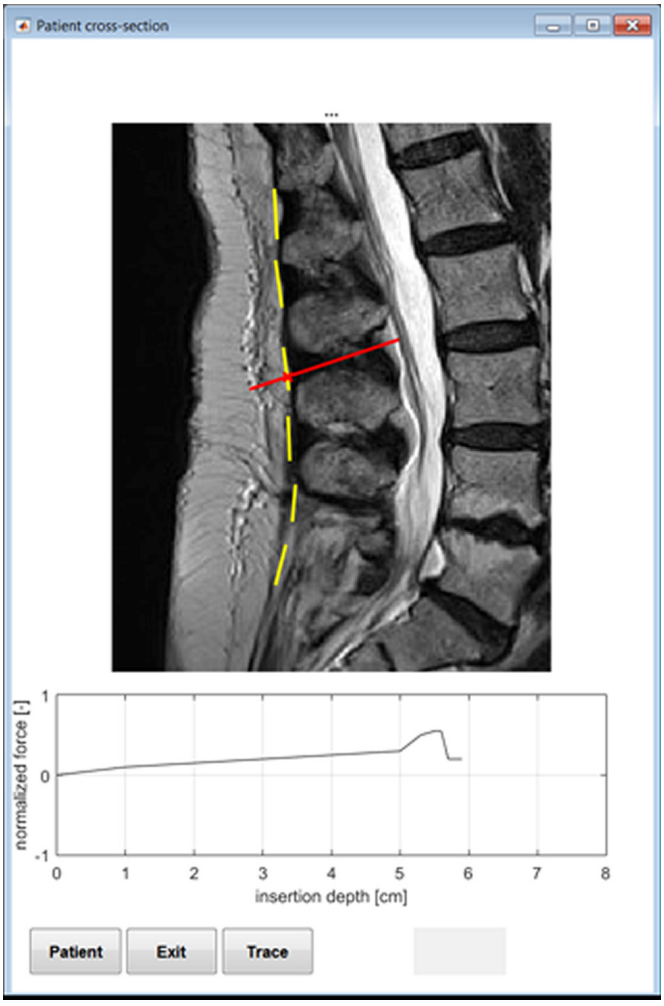

Figure 3 Screenshot of monitor visualizing real-time needle advancement during the procedures with visualization aid.

Those questions addressed participants' experience with the simulator and its advantage and added value as a teaching device. Participants were asked to provide their age, sex, and experience with epidural needle insertions. We registered the number of passes, bone contacts, dura contacts, spinal taps, as well as the time for the epidural procedure. The participants received feedback on their performance after completion of the study task and after answering all questions on the form.

All data were analyzed using IBM SPSS Statistics for Windows V.22.0. We tested the simulator for face and construct validity. Face validity was tested by assessing the "realism" of the simulator based on the feedback of the participants, who rated their experience on a Likert scale (strongly disagree=1, strongly agree=5). The consistency of the questionnaire was tested with Cronbach's alpha score. The simulator's construct validity was evaluated by comparing experienced and inexperienced groups for bone contacts, dura contacts, spinal taps, time taken for epidural needle placement, and number of attempts. The correlations were tested by Pearson's χ^2 test. The influence of the visual aid was assessed by comparing the results with the visualization aid on or off by means of the Wilcoxon signed-rank test. The effect of visualization aid prior to performance without visualization was tested by Mann-Whitney U test. A p value of <0.05 was considered statistically significant.

Table 1 Demographics

Group	Novice	Experts	Total
Total number (male/female)	20 (11/9)	48 (32/16)	68 (25/43)
Median age (range)	29 (23–35)	37 (29–66)	35 (23–66)

RESULTS

Sixty-eight participants were included in the study. The participants were divided into two groups based on their previous experience in epidural needle placement. Forty-eight participants were defined as "expert group" (more than 30 epidural needle placements) and 20 participants were assigned to the "novice group". Demographic data are displayed in table 1.

The face validity questionnaire revealed a satisfactory overall score of 3.7 (2.0–4.8) on a 5-point scale. The highest scores were given for the usefulness of the simulator (4.15±0.83) for hand-eye coordination and intuitivity, while the lowest scores were given for the adequacy of the simulator to measure performance and ligamentum flavum resistance (table 2). High marks were also given for the LOR experience, with experts grading 4.0±0.9 and novice 3.8±0.9 on average. Scores regarding face validity given by experts and novices did not differ significantly. For the questionnaire's consistency and reliability, a Cronbach's alpha score of 0.82 was calculated.

Table 3 illustrates no significant difference between the experienced and novice groups regarding total bone contacts, number of attempts, and procedure time, although there was a slight tendency for an increased total number of attempts in the novice group (p=0.06). However, the novices had significantly more dura contacts (p=0.001) and spinal taps (p=0.04).

With the visual aid turned on, participants made significantly less bone contacts (2.8 [0–14] vs 4.3 [0–23]), needed less attempts (8.9 [6–19] vs 10.9 [6–29]), and required less time to finish the task (197 s [55–1079] vs 233 [53–605]) when compared with attempts without the visual aid. However, turning visual aid on or off made no statistically significant difference with respect to dura contacts and spinal taps (table 4).

Visualization in the first round led to less attempts and a shorter procedure time (table 5). There was no difference for bone contacts, dura contacts, and spinal taps between the groups that practiced with the visual aid first compared with those using visual support in the second round.

DISCUSSION

The face validity of our simulator is rated as good but not perfect (3.7 out of 5). In the clinically important measures, number of dura contacts or spinal taps, the experienced anesthesiologist performed significantly better than the novices. On the other hand, we found no differences regarding surrogate parameters as the number of attempts or bone contacts or total time required. Possibly experienced anesthesiologists were more cautious in the proximity of the epidural space. Turning on the visual aid decreased bone contacts, led to fewer attempts, and less time required. The difference in the number of dura contacts or spinal taps did not reach significance. Finally, practicing with visualization improved the performance time and decreased the number of attempts even after the visualization support was turned off. Although the results generally underline the validity of the simulator and the advantage of visualization, they also raise questions regarding adequate variables to measure good performance.

Overall satisfaction of the participants with the simulator was reasonable to good depending on the item asked. All 11 items were rated as good on average (>2.5 of 5 on a Likert scale),

Table 2 Rating of the simulator by participants

		Novice		Experts		Total	
	Question	Mean	SD	Mean	SD	Mean	SD
1	The simulator handling is intuitive.	4.2	1.06	3.96	0.71	4.03	0.83
2	The resistance of the plunger is realistic.	3.55	0.76	3.63	1.04	3.60	0.96
3	Movement of the needle through the simulated tissue is realistic.	3.5	0.61	3.15	0.99	3.25	0.9
4	Bone contact with the needle is realistic.	3.45	1.05	3.79	0.77	3.69	0.87
5	The resistance of the ligamentum flavum is realistic.	3.05	0.69	3.23	0.93	3.18	0.86
6	The loss of resistance is realistic.	3.75	0.91	4.04	0.92	3.96	0.92
7	The simulator looks realistic.	3.6	0.6	3.29	1.07	3.38	0.96
8	The simulator is useful for training hand-eye coordination.	4.4	0.82	4.04	0.82	4.15	0.83
9	The simulator is an adequate tool to learn placing epidural catheters.	4.2	0.62	3.9	0.99	3.99	0.91
10	The simulator is a useful tool for educating anesthesiologists.	4.15	0.81	3.88	0.94	3.96	0.9
11	The simulator is an adequate tool for measuring performance.	3.15	0.81	2.88	1.18	2.96	1.08

with very good ratings for usefulness for training hand-eye coordination (4.2) and intuitive handling (4.0). On the other hand, the simulation of the ligamentum flavum and movement of the needle through the tissues, as well as the appearance of the simulator, were rated less favorably. Since resistance to needle movement is not influenced by faster or slower movement of the needle, the handling of the needle feels rather unnatural and might explain the lower marks given on this parameter. Having only 2 degrees of freedom, our simulator implemented some, but not all, features of the real procedure, and therefore participants agreed less on the statement that the simulator can be used to measure performance.

There is no unique variable to measure performance of procedures in regional anesthesia. Usually, the ability to perform a block under experienced supervision without help is rendered as "success" in clinical studies. However, more objective measures are not validated. Therefore we used five different variables to measure success. We defined the avoidance of dura contact and spinal tap as being clinically most important. Since we expected the incidence to be low, we also instituted three surrogate variables (number of bone contacts, number of attempts, and time required). Clinical experience was significantly correlated with less dura contacts and spinal taps. Surprisingly this was not the case with other surrogate parameters (table 3). Two other studies using different simulators also failed to demonstrate a correlation between previous experience and bone contacts, procedure time, and number of attempts.[5 16] In our study, procedure time was not correlated at all, but there was a (not significant) tendency toward less number of attempts and bone contacts by the experienced group. It seems as if experience becomes important during the more crucial part of the procedure. However, this is just one possible interpretation, and it might also be possible that the simulator was more realistic when the epidural space was reached.

Compared with clinical practice, all participants seemed to require a large amount of time, and had more attempts and bone contacts. This may be due to the fact that the simulator was based on tomographic pictures of the lumbar spine taken in the supine position. In the clinical situation, the epidural puncture is performed on a flexed vertebral column in the sitting or lateral decubitus position, causing the opening of the posterior interlaminar space and thus changing the relationship of the osseous and soft tissues.[17 18] This might be the reason for the relatively high number of bone contacts and attempts in both groups. Furthermore, this might have led to equality between groups. However, after reaching the ligamentum flavum, the situation seemed to be more realistic and here the performance of experienced group was superior. Thus, regarding the clinically important measures, construct validity was demonstrated, whereas it remains unclear why this did not show up in the less important surrogate parameters.

As expected, after enabling the visual aid, participants made less bone contacts, needed fewer attempts, and required less time to finish the task. This is in accordance with data proving the advantage of prescan ultrasound imaging on the success rate of epidural punctures.[19 20] Mirroring the clinical situation, a prescan of the anatomic structures could improve the precision of the simulated puncture. The possible programming of our simulator with different radiologic scans could help future students: first take an ultrasound scan of the patient, upload this into the simulator, and practice this specific patient on the simulator before returning to the patient to do the procedure. Such an individualized planning may facilitate or enable otherwise difficult or impossible punctures.

The impact of visual aid was also observed when participants who had it turned on for their first six attempts then performed the following six attempts without visual support. Hence, it

Table 3 Effect of experience on different parameters

Parameter	Mean (range)		P value
	Experienced	Novice	
Total number of attempts	18.7 (12–42)	22.1 (14–41)	0.06
Total bone contacts	6.2 (0–25)	8.7 (2–27)	0.10
Total procedure time (s)	425 (150–1684)	440 (207–708)	0.80
Total dura contacts	0.4 (0–3)	1.5 (0–6)	0.001
Total spinal taps	0.07 (0–2)	0.4 (0–4)	0.04

Table 4 Comparison between participants' performance with visual aid usage off or on

	Mean (range)		
Parameter	Visual aid off	Visual aid on	P value
Total number of attempts	10.9 (6–29)	8.9 (6–19)	0.01
Total bone contacts	4.3 (0–23)	2.8 (0–14)	0.02
Total procedure time (s)	230 (27–605)	194 (0–1079	0.02
Total dura contacts	0.4 (0–5)	0.3 (0–2)	0.07
Total spinal taps	0.12 (0–4)	0.04 (0–2)	0.4

Table 5 Comparison between participants' "blind" performance before or after training with the visualization aid

Parameter	Mean (range) Before visualization training, n=36	After visualization training, n=32	P value
Total number of attempts	11.9 (7–27)	9.7 (6–29)	0.047
Total bone contacts	4.9 (0–23)	3.5 (0–23)	
Total procedure time (s)	279 (70–574)	182 (53–605)	0.01
Total dura contacts	0.5 (0–5)	0.3 (0–2)	
Total spinal taps	0.2 (0–4)	0 (0)	

required less time and participants needed fewer attempts, which could be attributed to learning and acquiring proficiency and the benefit of the visual aid as a learning tool. However, the incidence of dura contact and spinal taps remained unaltered.

We could demonstrate face and construct validity of this simulator with only 2 degrees of freedom. Thus, even a low-fidelity simulator is useful in learning epidural punctures. However, we still have a long way to go before we develop a more realistic simulator with more degrees of freedom, more realistic feeling while advancing the needle, incorporation of ultrasound prescans into the simulator in order to individualize training, and finally the proof that novices could accelerate their learning curve using a simulator, having the "expert" skills when performing their first epidural on a real patient.

In conclusion, the TU-Delft SENS has a sufficient face and construct validity for teaching epidural needle placement to anesthesiology residents. We showed the value of real-time visualization and demonstrated that preprocedure visualization led to a higher precision. This was present even when the following simulations were done without visualization. Development of high-fidelity simulators for epidural punctures based on ultrasound prescans might abandon the need to train motor skills on a patient, and will enable or at least facilitate epidural punctures in anatomically difficult situations.

Correction notice This article has been corrected since it first published online. The open access licence type has been amended.

Funding Institutional funding.

Competing interests None declared.

Patient consent for publication Not required.

Provenance and peer review Not commissioned; externally peer reviewed.

REFERENCES

1 Friedman Z, Katznelson R, Devito I, *et al*. Objective assessment of manual skills and proficiency in performing epidural anesthesia--video-assisted validation. *Reg Anesth Pain Med* 2006;31:304–10.
2 Vaughan N, Dubey VN, Wee MY, *et al*. A review of epidural simulators: where are we today? *Med Eng Phys* 2013;35:1235–50.
3 Pedersen TH, Meuli J, Plazikowski EJ, *et al*. Loss of resistance: a randomised controlled trial assessing four low-fidelity epidural puncture simulators. *Eur J Anaesthesiol* 2017;34:602–8.
4 Carassiti M, Mattei A, Quarta R, *et al*. A new pressure guided management tool for epidural space detection: feasibility assessment on a simulator. *Artif Organs* 2017;41:E320–E325.
5 Friedman Z, Siddiqui N, Katznelson R, *et al*. Clinical impact of epidural anesthesia simulation on short- and long-term learning curve: high- versus low-fidelity model training. *Reg Anesth Pain Med* 2009;34:229–32.
6 Leighton BL. A greengrocer's model of the epidural space. *Anesthesiology* 1989;70:368–9.
7 Brett PN, Harrison AJ, Thomas TA. Schemes for the identification of tissue types and boundaries at the tool point for surgical needles. *IEEE Trans. Inform. Technol. Biomed.* 2000;4:30–6.
8 Brett PN, Parker TJ, Harrison AJ, *et al*. Simulation of resistance forces acting on surgical needles. *Proc Inst Mech Eng H* 1997;211:335–47.
9 Boessenkool W. Force and pressure feedback during epidural needle insertion in the ligamentum flavum of piglets. *Master Thesis in Biomedical Engineering published* 2012.
10 Manoharan V, van Gerwen D, van den Dobbelsteen JJ. Design and validation of an epidural needle insertion simulator with haptic feedback for training resident anaesthesiologists. *HAPTICS. IEEE* 2012:341–8.
11 Vogt M, van Gerwen DJ, Lubbers W, *et al*. Optimal point of insertion and needle angle in neuraxial blockade using a midline approach: a study in computed tomography scans of adult patients. *Reg Anesth Pain Med* 2017;42:600–8.
12 Konrad C, Schüpfer G, Wietlisbach M, *et al*. Learning manual skills in anesthesiology: is there a recommended number of cases for anesthetic procedures? *Anesth Analg* 1998;86:635–9.
13 Kopacz DJ, Neal JM, Pollock JE. The regional anesthesia "learning curve". What is the minimum number of epidural and spinal blocks to reach consistency? *Reg Anesth* 1996;21:182–90.
14 de Oliveira Filho GR. The construction of learning curves for basic skills in anesthetic procedures: an application for the cumulative sum method. *Anesth Analg* 2002;95:411–6.
15 Schupfer GK, Konrad C. Poelaert JI manual skills in anaesthesiology. *Anaesthesist* 2003;52:527–34.
16 Lee RA, van Zundert TC, van Koesveld JJ, *et al*. Evaluation of the Mediseus epidural simulator. *Anaesth Intensive Care* 2012;40:311–8.
17 Lee RA, van Zundert AA, Botha CP, *et al*. The anatomy of the thoracic spinal canal in different postures: a magnetic resonance imaging investigation. *Reg Anesth Pain Med* 2010;35:364–9.
18 Hermanides J, Hollmann MW, Stevens MF, *et al*. Failed epidural: causes and management. *Br J Anaesth* 2012;109:144–54.
19 Grau T, Leipold RW, Conradi R, *et al*. Efficacy of ultrasound imaging in obstetric epidural anesthesia. *J Clin Anesth* 2002;14:169–75.
20 Perlas A, Chaparro LE, Chin KJ. Lumbar neuraxial ultrasound for spinal and epidural anesthesia: a systematic review and meta-analysis. *Reg Anesth Pain Med* 2016;41:251–60.