A randomized comparison of video demonstration versus hands-on training of medical students for vacuum delivery using Objective Structured Assessment of Technical Skills (OSATS)

Ziad Hilal, MD∗, Anne K. Kumpernatz, MD, Günther A. Rezniczek, PhD, Cem Cetin, MD, Eva-Katrin Tempfer-Bentz, MD, Clemens B. Tempfer, MD, MBA

Abstract

Background: To compare medical students’ skills for vaginal operative delivery by vacuum extraction (VE) after hands-on training versus video demonstration.

Methods: We randomized medical students to an expert demonstration (group 1) or a hands-on (group 2) training using a standardized VE algorithm on a pelvic training model. Students were tested with a 40-item Objective Structured Assessment of Technical Skills (OSATS) scoring system after training and 4 days later. OSATS scores were the primary outcome. Performance time, self-assessment, confidence, and global rating scale were secondary outcomes. We assessed the constructive validity of OSATS in this VE model comparing metric scores of experts and students.

Results: In all, 137 students were randomized. OSATS scores were higher in group 2 (n=63) compared with group 1 (n=74) (32.89±6.39 vs 27.51±10.27, respectively; P<0.0001). Global rating scale (1.49±0.76 vs 2.33±0.94, respectively; P<0.0001), confidence (2.22±0.75 vs 3.26±0.94, respectively; P=0.04), self-assessment (2.03±0.62 vs 2.51±0.77, respectively; P<0.0001), and performance time (38.81±11.58 seconds vs 47.23±17.35 seconds, respectively; P=0.001) also favored group 2. After 4 days, this effect persisted with OSATS scores still being significantly higher in group 2 (30.00±6.50 vs 25.59±6.09, respectively; P=0.001). The assessed OSATS scores showed constructive validity. In a multiple linear regression analysis, group assignment independently influenced OSATS scores, whereas sex, handedness, sports activities, and type of curriculum were not independently associated with OSATS scores.

Conclusions: Hands-on training is superior to video demonstration for teaching VE on a pelvic model.

Abbreviations: CON = confidence, GRS = global rating scale, OSATS = Objective Structured Assessment of Technical Skills, PT = performance time, SA = self-assessment, VE = vacuum extraction.

Keywords: delivery, operative vaginal delivery, OSATS, randomized controlled trial, technical skills, training, vacuum extraction

1. Introduction

Vaginal operative delivery occurs in 15% of all deliveries.[1] Forceps and vacuum extractors are commonly used for vaginal operative delivery based on randomized trials demonstrating the feasibility, safety, and comparable efficacy of these methods.[2] However, the use of those instruments can provoke serious complications such as fourth-degree tears, fetal facial injury, or fetal asphyxia.[3-6]

Whatever the specific technique used in a given obstetric hospital, teaching obstetric residents the technique of vacuum extraction (VE) is challenging. Traditionally, residents gain experience with VE performance using a learning-by-doing approach under expert guidance. However, many vaginal
operative deliveries are performed under stressful conditions due to maternal or fetal distress. This interferes with the efficiency of the teaching and learning process, since it is well-known that teaching and also learning performance is hampered by stress.[7] Also, it is ethically questionable to teach residents obstetric techniques with women as study objects, especially in the early stages of surgical training.[9]

Low and high-fidelity training models and simulators have therefore been tested as potentially effective tools to enhance technical performance of medical staff.[8–11] Based on these data, standardized trainings on pelvic models may be potentially useful to prepare residents before they start to perform VEs on women. No data have been published regarding the validity of various training methods of VE (PUBMED search; 03-february-2016; search terms: vacuum extraction, training, model, randomized, pelvic trainer, vaginal operative delivery).

Since surgical and manual obstetric techniques such as hysteroscopy, resolution of a shoulder dystocia, and vaginal delivery for breech presentation can be effectively trained with dummy models,[12–16] it is reasonable to assume that dummy trainings for VE may also be effective. Previously, our group has demonstrated that Objective Structured Assessment of Technical Skills (OSATS) in combination with pelvic dummy training can effectively improve the skills of obstetric trainees.[12,14] Specifically, we have shown in a randomized trial of 172 medical students that OSATS scores, performance time (PT), global effectiveness, and confidence (CON) for vaginal delivery for breech presentation were significantly higher after hands-on dummy training compared with demonstration alone.[14] In another randomized trial, we randomized 203 probands to a 30-minute hands-on and a 30-minute demonstration training session, teaching a standardized shoulder dystocia management scheme on a pelvic training model. In this study, hands-on dummy training helped to achieve a significantly higher effectiveness and confidence in OSATS scores, and also PT, global effectiveness, self-assessment (SA), and CON were significantly higher after hands-on dummy training.[12]

The present study adds to and builds on these previous OSATS studies in a number of ways. For example, in the present study, we have added participant characteristics such as handedness and sports activities to the multivariate analysis. Also, the OSATS procedure has been improved compared with the previous studies using a number of metric scores. Also, we used a standardized teaching video in the control group to avoid intra and interteacher variabilities observed in previous experiments. Lastly, we tested for constructive validity using a comparison of metric scores between experts and nonexperts.

In the present study, we tested the hypothesis that hands-on training of VE in a pelvic training model is more effective than expert video training with effectiveness defined as high OSATS scores. To test this hypothesis, we designed a randomized controlled trial measuring OSATS scores of students undergoing a defined VE management algorithm after a 30-minute hands-on training compared with a 30-minutes expert demonstration using a training video (https://youtu.be/XpZOHsQ2-qA). To test for immediate and also delayed training effects, we tested students immediately after training (day 1) and retested them after 4 days (day 3).

2. Methods

This randomized, single-blinded, single-center trial was carried out between May 2015 and October 2015 at the Department of Obstetrics and Gynecology, Ruhr University Bochum, Bochum, Germany. The study participants were medical students who did an Obstetrics rotationship. This randomized study was approved by the local Ethics Committee (approval number: 4999–14; Date of approval: 23.10.2014). The trial was registered with ClinicalTrials.gov (NCT02468427). Informed consent was obtained from all study participants. Figure 1 shows a CONSORT diagram of the students flow through the study.

There were no specific inclusion or exclusion criteria other than informed consent and language barrier. For group assignment, we used a computer-generated randomization list, with a 1:1 allocation ratio without blocking. After randomization, students were allocated to different rooms, where the hands-on training and the video training session took place. The assessor (A.K.), who tested the students afterwards, was not present at the allocation, during the hands-on training, and during the video training session. Thus, the assessor was unaware of the participants’ group assignment. Students allocated to group 1 underwent a 30-minute video training session, watching an instruction video demonstrating all 40 items of VE as listed in Table 1. The video was produced by our Department and is in use for resident training purposes. In this 30-minute training video, one of the authors (C.T.) acts as an expert instructor, demonstrating all 40 steps of performing a VE according to the OSATS checklist used in this study (Table 1). For the training video, we used the same pelvic training model also used in the study (Phantom; Schultes Medacta, Germany). Students allocated to group 2 underwent a 30-minute hands-on training session with the students performing all 40 items of VE as listed in Table 1 at least once under expert instruction. The key maneuvers were as follows: vaginal digital examination, correct diagnosis of fetal lie,
Establishes correct position of the fetus. Also, additional items such as establishing the indication for VE, information of a senior obstetrician and anesthesiologist, or addition of oxytocin and emptying of the bladder were part of the OSATS (Table 1). The pelvic model and the model of the baby used in this study allowed the correct determination of the fetal position. We used a soft cup, single use vacuum extractor (KIWI OmniCup Single Use VAC 6000M; Clinical Innovations LLC, South Murray), licensed in the EU, for obstetric use.

Students did not take photographs or cell phone videos during the study procedures and they were not allowed to take written notes. After the training session, study participants were tested by an assessor (A.K.). Testing was repeated 4 days later. The assessor was unaware of the group assignment of the study participants. Testing of the study participants was performed using the OSATS technique. Specifically, the assessor calculated an OSATS score by adding points given for each of the 40 items on a task-specific check list (Table 1), with 1 point for correctly performing each item and 0 point for not performing or not correctly performing the item. Thus, a higher score indicates greater proficiency. The secondary outcomes SA, CON, and global rating scale (GRS) were graded using a 5-scale Likert scale, with lower values denoting better performance.

We tested the construct validity of the OSATS setting used in this study by comparing predefined metric scores between 6 experts and the 137 medical students recruited for this study. The metric scores were: number of cup detachments, pressure applied by the vacuum extractor (measured in mm Hg using the pressure scale attached to the KIWI OmniCup vacuum extractor), PT, and OSATS scores. PT was defined as the time between the starting point (start of item 1 of the OSATS checklist) and the stop time (completion of item 40 of the OSATS checklist). The 6 experts were consultants and residents with at least 4 years of practical obstetrical experience. We defined construct validity as given, when all metric scores were different between the 2 groups favoring the experts.

We chose to randomize medical students, because they had no previous obstetrics training or practical obstetric experience. In contrast to residents with various degrees of expertise and personal experience, this fact makes students ideal study participants in the sense that they make training effects easily visible.

Categorical variables were analyzed by chi-square test, and continuous variables were compared using the Mann–Whitney U test, with a significance level of 0.05. We performed a multiple linear regression model to test whether the training effect, as measured by OSATS scores, was independent of potential confounders such as the students’ sex (male vs female) or handedness. In addition, we included regular sports activities as a potential confounder, based on the assumption that engagement in regular sports may select a population with expert manual and coordination skills. Values are given as medians or means. A power calculation demonstrated that, with a sample size of 137, the study has a power of >80% to detect an absolute 20% difference in OSATS scores at a significance level of 0.05 using a Mann–Whitney U test. An OSATS score difference of 20%, that is, a median score difference of 4, range 0 to 8, was estimated to be achievable by hands-on training and to be clinically relevant based on previous OSATS training trials.[13,16] We used the statistical software SPSS 11.0 for Windows (SPSS Inc., Chicago, IL) for statistical analysis.

3. Results

We randomized 137 students. Table 2 shows the characteristics of the study population. Table 3 shows the primary and secondary outcomes broken down by group assignment. Specifically, we found that the OSATS scores (the primary outcome) were significantly higher in group 2 (n = 63) compared with group 1 (n = 74) (32.89 ± 6.39 vs 27.51 ± 10.27, respectively; P < 0.0001). Figure 2 shows a box plot of OSATS scores in both groups. The secondary outcomes GRS (1.49 ± 0.76 vs 2.33 ± 0.94, respectively; P < 0.0001), CON (2.22 ± 0.75 vs 3.26 ± 0.94, respectively; P = 0.04), SA (2.03 ± 0.62 vs 2.51 ± 0.77, respectively; P < 0.0001), and PT (38.81 ± 11.58 vs 47.23 ± 17.33 seconds, respectively; P = 0.001) were also significantly different favoring group 2.

### Table 1

| Objective Structured Assessment of Technical Skills (OSATS) checklist using 40 items to test proficiency in performing a vacuum extraction (VE) on a pelvic training model. | Yes | No |
|---|---|---|
| Recognizes the following indications for VE: | | |
| 1. Prolonged second stage of labor | | |
| 2. Nonmeasuring fetal heart rate tracings | | |
| 3. Tired labor contractions | | |
| 4. Maternal noncompliance | | |
| 5. Calls senior physician | | |
| 6. Calls anesthesiologist | | |
| 7. Calls neonatologist | | |
| 8. Emptying of the bladder | | |
| 9. Checks/apples intravenous aditus | | |
| 10. Starts/increases oxytocin | | |
| 11. Checks fetal heart rate tracings | | |
| Establishes correct position of the fetus | | |
| 12. Height | | |
| 13. Lie | | |
| 14. Presentation | | |
| 15. Position | | |
| 16. Checks the functionality of the vacuum extractor | | |
| 17. Applies local anesthesia to the perineum | | |
| 18. Performs a mediolateral episiotomy | | |
| 19. Places the cup correctly | | |
| 20. Builds up vacuum (pressure measurement in mm Hg) | | |
| 21. Checks for vaginal wall entrapment by vacuum cup | | |
| 22. Performs test traction | | |
| 23. Traction synchronous with next contraction | | |
| 24. Traction in the correct direction to deliver fetal head | | |
| 25. Vacuum release after delivery of the head | | |
| 26. Removes vacuum extractor | | |
| 27. Takes fetal head with both hands biauricularly | | |
| 28. Downward traction of the head | | |
| 29. Delivers anterior shoulder | | |
| 30. Places fetal head on left hand | | |
| 31. Support of the perineum | | |
| 32. Delivers posterior shoulder | | |
| 33. Delivers fetal body | | |
| 34. Waits until cord pulsation has ceased | | |
| 35. Cord clamping | | |
| 36. Hands over fetus to neonatologist | | |
| 37. Delivery of the placent, checks completeness | | |
| 38. Episiotomy repair | | |
| 39. Checks firmness of the uterus | | |
| 40. Writes VE report | | |
| Total score | /40 | |
| Time needed to complete sec | | |
Group 2 had a significantly higher proportion of participants achieving 90%, 80%, and 70% of the maximum OSATS scores (3/74, 17/74, and 37/74 vs 27/63, 46/63, and 52/63; all comparisons $P < 0.0001$).

To test for a delayed training effect, we retested students after 4 days. Fifty-four students in group 1 and 44 students in group 2 took part in retesting. There was no significant difference in the lost to follow-up rate between groups 1 and 2 (20/74 vs 19/63; $P = 0.4$). After 4 days, OSATS scores were still significantly higher in group 2 (30.00 ± 6.50 vs 25.59 ± 6.09, respectively; $P = 0.001$). After 4 days, the secondary outcome parameters GRS, SA, CON, and PT were not significantly different between groups 2 and 1 (2.00 ± 0.92 vs 1.91 ± 0.88, respectively; $P = 0.6$ for GRS; 2.44 ± 0.72 vs 2.43 ± 0.64, respectively; $P = 0.9$ for SA; 2.57 ± 0.83 vs 2.68 ± 0.78, respectively; $P = 0.5$ for CON; 34.38 ± 15.41 vs 36.92 ± 11.99 seconds, respectively; $P = 0.4$ for PT).

Students differed regarding their sex (female; $n = 87$ vs male; $n = 50$), handedness (right-handed; $n = 129$ vs left-handed; $n = 8$), regular sports activities (defined as at least 3 times/wk; yes; $n = 71$ vs no; $n = 66$), and regarding the type of Medical School curriculum (regular curriculum; $n = 110$ vs an experimental “model” curriculum; $n = 27$). Assessing these potential confounders, we found that males and females achieved equally high OSATS scores (31.12 ± 12.04 vs 29.37 ± 6.87, respectively; $P = 0.2$). Right-handers and left-handers also performed equally (29.89 ± 9.23 vs 31.75 ± 4.30, respectively; $P = 0.5$), as did those who did or did not engage in regular sports activities (30.37 ± 10.61 vs 29.61 ± 7.13, respectively; $P = 0.6$). Lastly, students of the regular curriculum and students of the model curriculum also performed equally (30.26 ± 9.42 vs 28.96 ± 7.59, respectively; $P = 0.5$). In a multiple linear regression analysis, group assignment (odds ratio [OR] 7.72, 95% confidence interval [CI] 3.49–17.06, $P < 0.0001$) independently influenced OSATS scores, whereas sex, handedness, sports activities, and type of curriculum were not independently associated with OSATS scores. Therefore, group assignment was an independent predictor of OSATS scores.

The assessed OSATS scores showed constructive validity. Specifically, the metric scores were different between the experts ($n = 6$) and the nonexperts ($n = 137$), all favoring the expert group (number of cup detachments [0 vs 2, respectively], pressure applied by the vacuum extractor [0.78 ± 0.04 vs 0.76 ± 0.07 mm Hg, respectively], PT [33.83 ± 10.57 vs 43.68 ± 15.09 seconds, respectively], and OSATS scores [37.83 ± 0.75 vs 30.50 ± 10.79, respectively]).

### 4. Discussion

In the present study, we found that structured hands-on training on a pelvic dummy is more effective than a teaching video regarding OSATS scores for technical achievement, and also CON and procedural speed when performing a delivery by VE. In a multivariate linear regression model, group assignment was an independent predictor of OSATS scores. Based on these results, hands-on training helps to achieve a significant improvement of VE management skills compared with expert video demonstration, with the effect persisting for at least 5 days. Based on these
results, we conclude that hands-on training is superior to expert video demonstration and should be recommended for teaching residents how to perform VE on a pelvic model.

Surgical simulators and pelvic training models have been described as useful tools for residents and advanced physicians for the training of surgical skills. However, no data are available regarding the effectiveness of dummy trainings for VE delivery. In our study, we assessed training performance, which may not be consistent with the performance in a real clinical situation. Our study suggests, but does not prove, that better VE performance on a model will also result in better performance in women. This assumption is supported by literature demonstrating improved real-life performance of shoulder dystocia management, and also laparoscopic surgery after simulator training. Lastly, not all OSATS items used in our study may be the standard of care of VE in other institutions.

In summary, the data collected in this prospective, randomized study provide level I evidence that hands-on training helps to achieve a significant improvement of VE performance over expert demonstration in a pelvic training model. These data support the incorporation of hands-on training and OSATS for teaching and training of VE into educational curriculums.

5. Conclusions

Hands-on training is superior to video demonstration for teaching VE on a pelvic model.

Acknowledgment

We acknowledge support by the DFG Open Access Publication Funds of the Ruhr-Universität Bochum.

References

[1] Popowski T, Porcher R, Fort J, et al. Influence of ultrasound determination of fetal head position on mode of delivery: a pragmatic randomized trial. Ultrasound Obstet Gynecol 2015;46:520–5.
[2] Williams MC, Knappel RA, O’Brien WF, et al. A randomized comparison of assisted vaginal delivery by obstetric forceps and polycyhelent vacuum cup. Obstet Gynecol 1991;78:789–94.
[3] O’Mahony F, Hofmeyr GJ, Menon V. Choice of instruments for assisted vaginal delivery. Cochrane Database Syst Rev 2010;CD003455.
[4] Ismail NA, Saharan WS, Zaleha MA, et al. Kiwi Omnicup versus Malmstrom metal cup in vacuum assisted delivery: a randomized controlled trial of the Kiwi Omnicup versus conventional ventouse cups. BJOG 2006;113:183–9.
[5] Kuit JA, Eppinga HG, Wallenburg HC, et al. A randomized comparison of vacuum extraction delivery with a rigid and a pliable cup. Obstet Gynecol 1993;82:280–4.
[6] Groom KM, Jones BA, Miller N, et al. A prospective randomised controlled trial of the Kiwi Omnicup versus conventional ventouse cups for vacuum-assisted vaginal delivery. BJOG 2006;113:183–9.
[7] Roberts MJ, Gale TC, McGrath JS, et al. Rising to the challenge: acute stress appraisals and selection centre performance in applicants to postgraduate specialty training in anaesthesia. Adv Health Sci Educ Theory Pract 2016;21:323–39.
[8] Nagendran M, Gurusamy KS, Aggarwal R, et al. Virtual reality training for surgical trainees in laparoscopic surgery. Cochrane Database Syst Rev 2013;8:CD006575.
[9] Agha RA, Fowler AJ. The role and validity of surgical simulation. Int Surg 2015;100:350–7.
[10] Goff BA, VanBlaricom A, Mandel L, et al. Comparison of objective, structured assessment of technical skills with a virtual reality hysteroscopy trainer and standard latex hysteroscopy model. J Reprod Med 2007;52:407–12.
[11] ACOG Committee on Obstetric Practice. ACOG Committee Opinion No. 340. Mode of term singleton breech delivery. Obstet Gynecol 2006;108:235–7.
[12] Buerkle B, Pueth J, Heller LA, et al. Objective structured assessment of technical skills evaluation of theoretical compared with hands-on training of shoulder dystocia management: a randomized controlled trial. Obstet Gynecol 2012;120:809–14.
[13] Alici F, Buerkle B, Tempfer CB. Objective Structured Assessment of Technical Skills (OSATS) evaluation of hysteroscopy training: a prospective study. Eur J Obstet Gynecol Reprod Biol 2014;178:1–5.
[14] Buerkle B, Rueter K, Heller LA, et al. Objective Structured Assessment of Technical Skills (OSATS) evaluation of theoretical versus hands-on training of vaginal breech delivery management: a randomized trial. Eur J Obstet Gynecol Reprod Biol 2013;171:252–6.
[15] Draycott TJ, Crofts JF, Ash JP, et al. Improving neonatal outcome through practical shoulder dystocia training. Obstet Gynecol 2008;112:14–20.
[16] Lavin JP Jr, Eaton J, Hopkins M. Teaching vaginal breech delivery and external cephalic version: A survey of faculty attitudes. J Reprod Med 2000;45:808–12.
[17] Heller L, Grimm C, Kueronya V, et al. A novel training model for the loop electrosurgical excision procedure. Am J Obstet Gynecol 2012;206:1–4.
[18] Egi H, Tokunaga M, Hattori M, et al. Evaluating the correlation between the HUESAD and OSATS scores: concurrent validity study. Minim Invas Ther Allied Technol 2013;22:144–9.
[19] Steehler MK, Chu EE, Na H, et al. Teaching and assessing endoscopic sinus surgery skills on a validated low-cost task trainer. Laryngoscope 2013;123:841–4.
[20] Dhingra S, Raff F. Obstetric trainees’ experience in VBD and ECV in the UK. J Obstet Gynaecol 2010;30:10–2.
[21] Larsen CR, Soerensen JL, Grantcharow TP, et al. Effect of virtual reality training on laparoscopic surgery: randomised controlled trial. BMJ 2009;338:b1802.