AOGS
ACTA Obstetricia et Gynecologica Scandinavica

AOGS MAIN RESEARCH ARTICLE

# Vacuum extraction: development and test of a procedure-specific rating scale

MATHILDE MAAGAARD[1], JEANETT OESTERGAARD[1], MARIANNE JOHANSEN[1], LISE LOTTE ANDERSEN[2], CHARLOTTE RINGSTED[3], BENT OTTESEN[1] & JETTE L. SØRENSEN[1]

[1]*Juliane Marie Centre for Children, Women and Reproduction, Rigshospitalet, University Hospital of Copenhagen, Copenhagen,* [2]*Obstetric Department, Odense University Hospital, Odense, and* [3]*Centre for Clinical Education, Rigshospitalet, University Hospital of Copenhagen, Copenhagen, Denmark*

**Key words**
assessment, clinical competence, obstetrics, medical education, vacuum extraction

**Correspondence**
Mathilde Maagaard, Juliane Marie Centre for Children, Women and Reproduction, Rigshospitalet, University Hospital of Copenhagen, Obstetrics Department, Blegdamsvej 9, section 4074, 2100 Ø – Copenhagen, Denmark.
E-mail: mmn@rh.regionh.dk or mathildemaagaard@hotmail.com

**Conflict of interest**
The authors have stated explicitly that there are no conflicts of interest in connection with this article.

Please cite this article as: Maagaard M, Oestergaard J, Johansen M, Andersen LL, Ringsted C, Ottesen B, Sørensen JL. Vacuum extraction: development and test of a procedure-specific rating scale. Acta Obstet Gynecol Scand 2012;91:1453–1459.

Received: 10 April 2012
Accepted: 1 August 2012

DOI: 10.1111/j.1600-0412.2012.01526.x

## Abstract

*Objectives.* To develop and validate an Objective Structured Assessment of Technical Skills (OSATS) scale for vacuum extraction. *Design.* Two-part study design: Primarily, development of a procedure-specific checklist for vacuum extraction. Hereafter, validation of the developed OSATS scale for vacuum extraction in a prospective observational study. *Setting*. Rigshospitalet, University Hospital of Copenhagen. *Population.* For development, an obstetric expert from each labor ward in Denmark (28 departments) was invited to participate. For validation, nine first-year residents and 10 chief physicians with daily work in the obstetric field were tested. *Methods.* The Delphi method was used for development of the scale. In a simulated vacuum extraction scenario, first-year residents and obstetric chief physicians were rated using the developed OSATS scale for vacuum extraction to test construct validity of the scale. *Main outcome measures.* Consensus for the content of the scale. To test the scale of Cronbach's alpha, interclass correlation and differential item function was calculated in the prospective study. *Results.* 89% completed the first and 61% completed the second Delphi round. Hereafter, consensus was obtained. There was a significant difference between residents' and experts' performance for total score and for the score of the separate parts of the scale. Cronbach's alpha for total score and for the separate parts of the scale was 0.91–0.95 and interclass correlation 0.84–0.9. *Conclusions.* The OSATS scale for vacuum extraction is a reliable test for differentiating between competence levels in a simulated setting.

**Abbreviations:** OSATS, Objective Structured Assessment of Technical Skills; PROMPT, PRactical Obstetric Multi-Professional Training.

## Introduction

Throughout the past two decades there has been increasing focus on measuring doctors' clinical performance. Consequently, a need for reliable and validated assessment instruments has emerged (1,2).

An uncomplicated delivery can quickly change to an emergency situation requiring highly skilled staff. Vacuum extraction is the most frequently used method for assisted vaginal

### Key Message

An OSATS scale for assessment of physician's performance in vacuum extraction was developed and was shown to be a reliable test for differentiating between competence levels in a simulated setting.

**Global-rating part - general skills**

| | | | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|---|---|
| | A | **Economy of movements** | Many unnecessary movements | | Efficient motion but some unnecessary movements | | Maximum economy of movements |
| | B | **Confidence of movements: instrument handling** | Repeatedly makes tentative or awkward moves with instruments | | Competent use of instruments although occasionally appeared stiff or awkward | | Fluent movements with instruments and no awkwardness |
| | C | **Economy of time** | Too long time used to perform sufficiently | | Intermediate time used to perform sufficiently | | Minimal time used to perform sufficiently |
| | D | **Errors: respect for tissue (perineum, vulva and vaginal tissue)** | Frequently used unnecessary force on tissue or risk of damage by inappropriate use of instruments | | Careful handling of tissue but occasionally risk of (minimal) damage | | Consistently and appropriated handled of tissues with no risk of damage |
| | E | **Flow of operation/operative technique** | Imprecise, wrong technique in approaching the operative intervention | | Careful technique with occasional errors | | Fluent, safe and correct technique in all stages of the operative procedure |

**Figure 1.** The Objective Structured Assessment of Technical Skills (OSATS) scale for vacuum extraction.

delivery in industrialized countries (8—13% of all deliveries) (3–5). In 1998, the US Food and Drug Administration issued a public health notice: "Need for CAUTION When Using Vacuum Assisted Delivery Devices" (6) due to the higher mortality and morbidity observed among children delivered after vacuum extractions. One of the Food and Drug Administration recommendations was: "those persons using vacuum devices must be versed in the use". Hence, there is a need for valid assessment instruments to ensure performance standards and develop effective training programs.

The aim of this study was to develop and explore the reliability, content and construct validity of a procedure-specific rating scale for vacuum extraction. Content validity refers to the extent to which a scale represents all the important facets of the procedure it is going to measure. Construct validity refers to whether this assessment instrument can differentiate between supposedly different levels of competency.

## Material and methods

Inspired by the Objective Structured Assessment of Technical Skills (OSATS) for surgical residents scale (7), a procedure-specific rating scale for vacuum extraction was developed. The OSATS scale for vacuum extraction consisted of: (1) a global-rating part including five items (item A–E) and (2) a procedure-specific part consisting of 16 items (items 1–16). All items of the OSATS for vacuum extraction scale (global and procedure-specific) were rated on a 5-point scale where 1, 3 and 5 points were anchored with a behavioral description. The scale is presented in Figure 1.

The global-rating part of the OSATS scale for vacuum extraction was built on the original OSATS scale by Reznick et al. from 1997 except for the two items: "*use of assistants*" and "*knowledge of specific procedure*". These were excluded since they were represented in more detail in the procedure-specific part of the revised scale. The remaining items were modified for general performance assessment in vacuum extraction.

The procedure-specific part of the OSATS scale for vacuum extraction was based on a nationally recognized in-training assessment checklist for residents (8). To improve face and content validity this checklist was further developed with consensus-seeking involvement of chief physician representatives from all obstetrical departments in Denmark using a modified Delphi technique. The Delphi technique is an anonymous process where responses are collected and analyzed until consensus is achieved (9). Items may be excluded or added. Each item from the national in-training assessment checklist was rated from 1 to 5 points concerning its relevance and ability to assess a resident's performance in vacuum

## Procedure-specific part – specific skills

| | | | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|---|---|
| ASSESSMENT | 1 | **Abdominal examination - focus on 4. grip** | Not executed | | Executed but in an unstructured and incompetent manner | | Executed in a structured and competent manner |
| | 2 | **Vaginal examination for presenting part, rotation and station** | Not executed | | Executed but too quick and in an incompetent manner | | Executed in a structured and competent manner with appropriate conclusion |
| | 3 | **Relevant assessment of the clinical situation and previous obstetric history** | No information obtained | | Obtaining of some obstetric history but with lack of relevant informations | | Obtains a full and relevant history in a timely manner |
| | 4 | **Make sure fetus is monitored by CTG** | Pays no attention | | Pays attention but without appropriate reassurance of the quality of the CTG | | Pays attention and makes sure fetal monitoring is sufficient during the procedure |
| DRUGS | 5 | **Assessment of the need for Oxytocin** | Pays no attention to contractions | | Pays attention to contractions but without outspoken reflection or actions taken | | Pays attention and acts relevant |
| | 6 | **Consider analgesia and acts accordingly** | Pays no attention | | Pays some degree of attention but does not act accordingly | | Pays attention and acts accordingly |
| INFO | 7 | **Informs patient and partner about the procedure and the risk of complications for mother and child** | Does not inform patient and partner | | Does inform patient and partner but in an insufficient or inappropriate way | | Does inform the patient and partner in an understandable and sufficient way that appears relevant for the situation |
| TECH | 8 | **Checks the vacuum system and manages to assemble the cup** | Does not check | | Checks but in an insufficient manner | | Checks and manages to sample the cup |
| | 9 | **Applies the cup on the flexion point** | Pays no attention - applies the cup randomly | | Pays attention but applies the cup randomly | | Pays attention and applies the cup competently |

**Figure 1.** Continued.

extraction. Furthermore, suggestions for additional items were encouraged. An item was excluded if more than one chief physician rated it below 3. A new item was included if recommended by more than 40%. The anchoring of items, both in the global and the procedure-specific part of the OSATS scale for vacuum extraction, was developed as a collaborative effort between the main author and three obstetric experts and co-authors (J.L.S., M.J. and L.A.).

First-year residents and chief physicians were tested in a simulated vacuum extraction scenario using the PRactical Obstetric Multi-Professional Training (PROMPT) birthing simulator (10) and rated by the OSATS scale for vacuum extraction for testing construct validity. The high competence level group (experts) was recruited randomly from different departments and all worked daily in the obstetric field. The low competence level group (residents) was recruited

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| N I C A L | 10 | **Ensures that no maternal vaginal tissue is stuck between the fetal head and the cup** | Does not check and begins traction immediately | | Does check but moves hands in an incompetent manner | | Does check and hands are moved with competence |
| | 11 | **Traction during contraction and when the patient is pushing** | Tractions occur randomly without paying notice to contractions | | Tractions are carried out during contractions but without co-operation with the mother | | Tractions are carried out during contraction and when the mother is pushing |
| | 12 | **Direction of traction follows the pelvic curve** | Tractions are random and/or by pumping up and down | | Tractions are in the right direction but with pumping movements | | Continuous and smooth tractions that follows the pelvic curve |
| | 13 | **Consider the need for episiotomy** | Pays no attention | | Pays some attention but without drawing any conclusions | | Pays attention and acts in a relevant way |
| | 14 | **Protection of the perineum** | Pays no attention | | Pays attention but without putting supportive hand in the right position and/or not doing so in timely manner | | Putting the supportive hand in the right position in a timely manner |
| S A F E T Y | 15 | **Using safe parameters for maximum numbers of tractions** | No awareness | | Number of tractions is 5 or more | | Number of tractions is 3 or less |
| | 16 | **Safe parameters for maximum spent time** | Pays no attention | | Pays some attention but too late in the procedure | | Pays attention to time. Acts relevant after 10 and 15 min and stops after 20 min |

**Figure 1.** Continued.

after attending their first off-site mandatory course during the first year of residency and they had to have done fewer than five supervised vacuum extractions. None of the participants had previous experience with the PROMPT birthing simulator. Furthermore, the participants were unaware of which simulated clinical scenario they were going to participate in. Participation was voluntary and no patients were involved; therefore according to Danish regulations, ethical approval was not required.

Each participant was presented with the following scenario: "You are called to the delivery room to attend a 30-year-old nulliparous, who has been actively pushing (second stage of labor) for one hour and twenty minutes. She is exhausted and in despair. Cardiotocography is classified as reassuring (baseline 140/minute, normal variability, no complicated decelerations and accelerations are present). The fetus's head is on the pelvic floor, in an occipito-anterior position. During contractions there is no progression. Contractions are regular (every second minute) and good. The parturient is requesting your assistance. In the delivery room is an auxiliary nurse who is experienced but needs your specific instructions for assistance". The following equipment was available: a timer, delivery kit, two types of vacuum extractors [soft (silicon) and hard Malmström (metal)], analgesia and scissors for episiotomy if needed.

The participants were told to proceed with the delivery, as they would in a real clinical setting, explaining their thoughts and actions aloud. The participants were asked to stop the delivery once the head was born.

Each simulated scenario was video-recorded. Two independent raters using the OSATS scale for vacuum extraction assessed the recordings. The raters were selected based on their expertise within obstetrics and were employed at two different university hospital departments. The raters were blinded to the level of clinical experience of the participant. To familiarize the raters with the OSATS scale for vacuum extraction and ensure rater consistency, the raters jointly rated three test scenarios. Each of the raters individually assessed the three scenarios and afterwards they discussed their ratings to reach consensus. In Figure 2 a flow chart presents the steps used in validating the OSATS scale for vacuum extraction in a simulated setting.

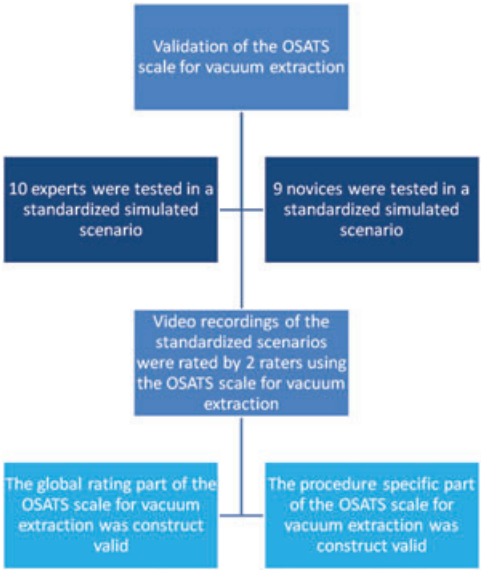


**Figure 2.** Validation of the Objective Structured Assessment of Technical Skills (OSATS) scale for vacuum extraction.

### *Statistical analyses*

Data were processed using SPSS 18.0 for Windows. Due to the sample size and nature of the results, a Gaussian distribution was not expected; therefore a non-parametric statistical test (Mann–Whitney U) was used to investigate differences between the groups. A $p$-value $< 0.05$ was considered significant. Scores are presented as medians and quartiles. Interclass-correlation and Cronbach's alpha were calculated separately for the total score, the global-rating part, and the procedure-specific part. Differential item function was calculated by logistic regression.

## Results

Figure 3 presents a flow chart of the development of the procedure-specific part of the OSATS scale for vacuum extraction. All invited residents and experts entered the study. The expert group consisted of 10 experts, five men and five women, from four different hospitals. Each expert had performed more than 100 vacuum extractions. The resident group consisted of nine first-year residents. The residents were all females and were employed in six different departments; they had been responsible for a vacuum-assisted delivery fewer than two times (range from 0–4) and none had seen the procedure being performed by a more experienced doctor more than 10 times. They had all previously performed the procedure on a mannequin.

Figure 4 shows how the total score and the score for each separate part of the OSATS scale for vacuum extraction can significantly differentiate between residents' and experts' performance levels.

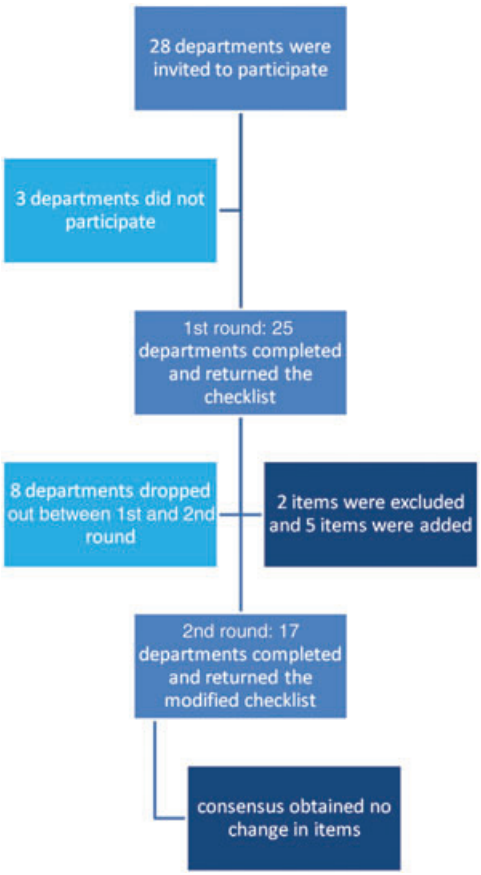


**Figure 3.** Validation of the in-training assessment checklist for residents by the Delphi method.

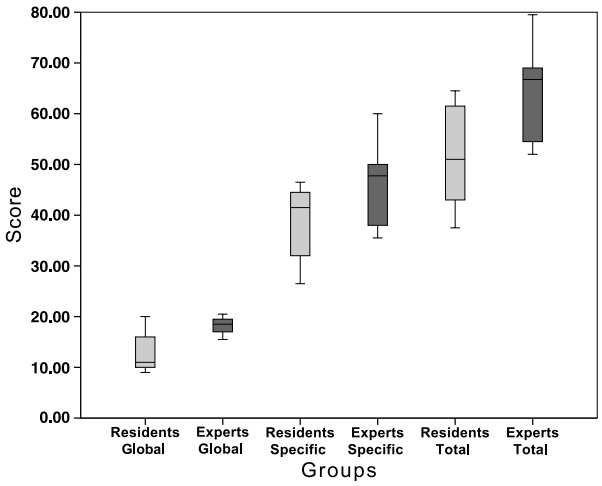


**Figure 4.** Box-plot of the score of the global-rating part, the procedure-specific part and total score of the Objective Structured Assessment of Technical Skills (OSATS) scale for vacuum extraction for residents and expert physicians in a simulated scenario. Global median score: 11 (residents) and 18.5 (experts), $p = 0.006$. Specific median score 39.5 (residents) and 46.25 (experts), $p = 0.037$. Total median score 46 (residents) and 63.25 (experts), $p = 0.01$.

The interclass-correlation coefficient based on the total score (all items), the global-rating part (items A–E) and the procedure-specific part (items 1–16) was 0.90, 0.85 and 0.84, respectively. Cronbach's alpha for the same parts was 0.95, 0.92 and 0.91, respectively.

Differential item function analysis using logistic regression was not useful in two of the 21 items, which could be explained by the size of this material. These were the items concerning safety parameters (numbers 14 and 15 in Figure 1). No differential item function was found in the remaining items, suggesting that there is no reason to exclude any items from the OSATS scale for vacuum extraction (11,12).

## Discussion

In this study an OSATS scale for vacuum extraction was developed and tested. The OSATS scale for vacuum extraction was found reliable for differentiating between competence levels in a simulated setting. The scale was based on an existing national in-training assessment checklist for residents, which was further developed through a consensus-seeking Delphi method process to ensure the content validity of this.

This study is the first to present a validated performance assessment instrument for vacuum extraction in a simulated environment. The results of this study are consistent with other studies: both the original study of the OSATS scale (7) and studies using the OSATS or a modified version of it in simulated or clinical settings (13–18). The original OSATS scale consisted of a global-rating scale with seven items rated on a 5-point scale (7). Reznick et al. have since published several studies with use of an OSATS scale and today they base the performance assessment on a combination of a global-rating part for general skills and a checklist part for procedure-specific skills (19–22). The global-rating part is always rated on a 5-point scale, whereas the checklist part is mostly rated on a dichotomous scale. However, the literature is versatile and many different ways of interpreting the OSATS scale can be found (15,19–25). We believe that the checklist part of the OSATS scale should be rated on a 5-point scale, where 1, 3 and 5 points are anchored with a behavioral description rather than on a dichotomy scale, as the former seems to be the best base for a clarified and elaborated feedback. For the same reason we chose to call this part of the OSATS scale for vacuum extraction the procedure-specific part instead of the checklist part. No matter how the scale is constructed, it is crucial to validate it in the context in which it is intended to be used.

We found the global-rating part to be superior in differentiating the level of competency, whereas the procedure-specific part is more complex in nature and will probably be of benefit as a guide for feedback to the individual physician. Our results also showed that, depending on the situation, each part of the scale can be used separately.

A disadvantage of the OSATS scale for vacuum extraction may be the complexity of many items. One way of overcoming this problem would be to rate video-recorded scenarios which gives the option of re-winding to improve the quality of the rating. Video-recorded simulated scenarios also allow assessment of performance and feedback to be delivered in a non-stressful setting. We validated the OSATS scale for vacuum extraction in a simulated environment where it is possible to create a standardized scenario that does not exist in real life. Therefore testing the OSATS scale for vacuum extraction in clinical settings requires subsequent testing of the same resident in several clinical situations due to inter-case variability (26).

Nonetheless, a valid performance assessment instrument is the first step towards developing a simulation-based training program in vacuum extraction. In a descriptive survey we found previously that most departments are performing simulation-based skills training in obstetric emergencies, but standardization of the training programs and assessment are still lacking to back this up (27).

This OSATS scale for vacuum extraction can be used for assessment of physician competence in a simulated setting before practicing on patients. Furthermore, it probably has potentials as a benchmark for structured elaborated feedback as part of a training program. Hopefully, through a validated training program we can facilitate better clinical practice and hence improve patient safety in the future. The OSATS scale for vacuum extraction is reliable for differentiating between competence levels in a simulated setting and has excellent inter-rater reliability and internal consistency.

## Funding

No special funding.

### References

1. Johannsson H, Ayida G, Sadler C. Faking it? Simulation in the training of obstetricians and gynaecologists. Curr Opin Obstet Gynecol. 2005;17:557–61.
2. Veltman LL. Getting to havarti: moving toward patient safety in obstetrics. Obstet Gynecol. 2007;110:1146–50.
3. Murphy DJ, Liebling RE, Verity L, Swingler R, Patel R. Early maternal and neonatal morbidity associated with operative delivery in second stage of labour: a cohort study. Lancet. 2001;358:1203–7.
4. O'Mahony F, Settatree R, Platt C, Johanson R. Review of singleton fetal and neonatal deaths associated with cranial trauma and cephalic delivery during a national intrapartum-related confidential enquiry. Br J Obstet Gynaecol. 2005;112:619–26.
5. Towner D, Castro MA, Eby-Wilkens E, Gilbert WM. Effect of mode of delivery in nulliparous women on neonatal intracranial injury. N Engl J Med. 1999;341:1709–14.

6. http://www.fda.gov/MedicalDevices/Safety/Alertsand Notices/PublicHealthNotifications/ucm062295.htm. 1998. Ref Type: Internet Communication.
7. Martin JA, Regehr G, Reznick R, MacRae H, Murnaghan J, Hutchison C, et al. Objective structured assessment of technical skill (OSATS) for surgical residents. Br J Surg. 1997;84:273–8.
8. http://www.dsog.dk/files/Checkliste%20gynobs%202007.pdf. 2011. Ref Type: Internet Communication.
9. Graham B, Regehr G, Wright JG. Delphi as a method to establish consensus for diagnostic criteria. J Clin Epidemiol. 2003;56:1150–6.
10. http://limbsandthings.com/global/products/prompt-birthing-simulator-standard/. 2011. Ref Type: Internet Communication.
11. Swaminathan H & Rogers HJ. Detecting differential item functioning using logistic regression procedures. J Educ Measurement. 1990;27:361–70.
12. Zumbo BD. Handbook on the Theory and Methods of Differential Item Functioning (DIF): Logistic Regression Modeling as a Unitary Framework for Binary and Likert-type (Ordinal) Item Scores. Ottawa: Department of National Defence, Directorate of Human Resources and Evaluation. 1999.
13. Datta V, Chang A, Mackay S, Darzi A. The relationship between motion analysis and surgical technical assessments. Am J Surg. 2002;184:70–3.
14. Datta V, Bann S, Mandalia M, Darzi A. The surgical efficiency score: a feasible, reliable, and valid method of skills assessment. Am J Surg. 2006;192:372–8.
15. Goff B, Mandel L, Lentz G, Vanblaricom A, Oelschlager AM, Lee D, et al. Assessment of resident surgical skills: is testing feasible? Am J Obstet Gynecol. 2005;192:1331–8.
16. Goff BA, Lentz GM, Lee D, Houmard B, Mandel LS. Development of an objective structured assessment of technical skills for obstetric and gynecology residents. Obstet Gynecol. 2000;96:146–50.
17. Larsen CR, Grantcharov T, Schouenborg L, Ottosen C, Soerensen JL, Ottesen B. Objective assessment of surgical competence in gynaecological laparoscopy: development and validation of a procedure-specific rating scale. Br J Obstet Gynaecol. 2008;115:908–16.
18. VanBlaricom AL, Goff BA, Chinn M, Icasiano MM, Nielsen P, Mandel L. A new curriculum for hysteroscopy training as demonstrated by an objective structured assessment of technical skills (OSATS). Am J Obstet Gynecol. 2005;193:1856–65.
19. Anastakis DJ, Wanzel KR, Brown MH, McIlroy JH, Hamstra SJ, Ali J, et al. Evaluating the effectiveness of a 2-year curriculum in a surgical skills center. Am J Surg. 2003;185:378–85.
20. Ault G, Reznick R, MacRae H, Leadbetter W, DaRosa D, Joehl R, et al. Exporting a technical skills evaluation technology to other sites. Am J Surg. 2001;182:254–6.
21. Dorman K, Satterthwaite L, Howard A, Woodrow S, Derbew M, Reznick R, et al. Addressing the severe shortage of health care providers in Ethiopia: bench model teaching of technical skills. Med Educ. 2009;43:621–7.
22. Schijven MP, Reznick RK, ten Cate OT, Grantcharov TP, Regehr G, Satterthwaite L, et al. Transatlantic comparison of the competence of surgeons at the start of their professional career. Br J Surg. 2010;97:443–9.
23. Collins AM, Ridgway PF, Hassan MS, Chou CW, Hill AD, Kneafsey B. Surgical instruction for general practitioners: how, who and how often? J Plast Reconstr Aesthet Surg. 2010;63:1156–62.
24. Setna Z, Jha V, Boursicot KA, Roberts TE. Evaluating the utility of workplace-based assessment tools for speciality training. Best Pract Res Clin Obstet Gynaecol. 2010;24:767–82.
25. Siddiqui NY, Stepp KJ, Lasch SJ, Mangel JM, Wu JM. Objective structured assessment of technical skills for repair of fourth-degree perineal lacerations. Am J Obstet Gynecol. 2008;199:676.
26. Konge L, Larsen KR, Clementsen P, Arendrup H, von BC, Ringsted C. Reliable and valid assessment of clinical bronchoscopy performance. Respiration. 2012;83:53–60.
27. Maagaard M, Johansen M, Lottrup P, Sorensen JL. Clinical skills training in obstetrics-a descriptive survey of current practice in Denmark. Acta Obstet Gynecol Scand. 2012;91:143–6.