

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/234823092>

# A Simple Framework for Assessing Technical Skills in a Resident Observed Structured Clinical Examination (OSCE): Vaginal Laceration Repair

ARTICLE in JOURNAL OF SURGICAL EDUCATION · JANUARY 2013

Impact Factor: 1.38 · DOI: 10.1016/j.jsurg.2012.08.005 · Source: PubMed

CITATIONS

3

READS

77

## 4 AUTHORS:



**Abigail Winkel**

NYU Langone Medical Center

8 PUBLICATIONS 8 CITATIONS

SEE PROFILE



**Veronica Lerner**

NYU Langone Medical Center

11 PUBLICATIONS 68 CITATIONS

SEE PROFILE



**Sondra Zabar**

NYU Langone Medical Center

69 PUBLICATIONS 560 CITATIONS

SEE PROFILE



**Demian Szyld**

NYU Langone Medical Center

30 PUBLICATIONS 155 CITATIONS

SEE PROFILE

# A Simple Framework for Assessing Technical Skills in a Resident Observed Structured Clinical Examination (OSCE): Vaginal Laceration Repair

Abigail Ford Winkel, MD, Veronica Lerner, MD, Sondra R. Zabar, MD, Demian Szyld, MD

Department of Obstetrics and Gynecology, New York University School of Medicine, New York, New York

**OBJECTIVES:** Educators of trainees in procedure-based specialties need focused assessment tools that are valid, objective, and assess technical skills in a realistic context. A framework for hybrid assessment using standardized patient scenarios and bench skills testing might facilitate evaluation of competency.

**METHODS:** Seven PGY-1 obstetrics and gynecology residents participated in a hybrid assessment that used observed structured clinical examination (OSCE) by a standardized patient who had sustained a vaginal laceration during vaginal delivery. The residents elicited a history and counseled the patient, and then completed a laceration repair on a pelvic model. The residents were rated on their performance in the scenario, which included issues of cultural competency, rapport-building, patient counseling. The technical skills were videotaped and rated using a modified global assessment form by 2 faculty members on a 3-point scale from “not done” to “partly done” to “well-done.” Residents also completed a subjective assessment of the station.

**RESULTS:** Mean technical performance of the residents on the technical skills was 55% “well-done,” with a range of 20%-90%. The assessment identified 3 residents as below the mean, and 1 resident with areas of deficiency. Subjective assessment by the residents was that juggling the technical, cognitive, and affective components of the examination was challenging.

**CONCLUSIONS:** Technical skills can be included in a case-based assessment using scenarios that address a range of cognitive and affective skills required of physicians. Results may help training programs assess individuals' abilities as well as identify program needs for curricular improvement. This framework might be useful in setting standards for competency and identifying poor performers. (J Surg 70:10-14. © 2012 Association of Program Directors in Surgery. Published by Elsevier Inc. All rights reserved.)

*Correspondence:* Inquiries to Abigail Ford Winkel, MD, Department of Obstetrics and Gynecology, New York University School of Medicine, 550 First Avenue, NBV 9E2, New York, NY 10016; e-mail: [abigail.winkel@nyumc.org](mailto:abigail.winkel@nyumc.org)

**KEY WORDS:** assessment, residency, technical skills, observed structured clinical examination, episiotomy, vaginal laceration repair

**COMPETENCIES:** Patient Care, Medical Knowledge, Professionalism, Interpersonal and Communication Skills, Practice-Based Learning and Improvement

## INTRODUCTION

Graduate medical educators have been called to create competency-based educational programs by the Accreditation Council for Graduate Medical Education (ACGME), Medicare and the public.<sup>1,2</sup> Focused assessment tools that are objective, valid and realistic are needed to assess residents' skills across a range of competencies.<sup>1</sup> Traditional assessment of trainees' by subjective performance evaluations can be affected by several biases, including the halo effect, recall bias, and insufficient exposure.<sup>3,4</sup> It is challenging for faculty to objectively assess the complete range of competencies using these traditional tools.<sup>5</sup> Surgical educators have developed several objective tools for formative and summative evaluation of resident technical skills.<sup>4-7</sup> Objective structured assessment of technical skills (OSATS) have been validated across a variety of procedures outside of the clinical environment, which allows for safe and structured assessment.<sup>4-9</sup> However, evaluating technical ability in a vacuum of the other skills a surgeon needs—communication skills, rapport-building, professionalism—gives an incomplete picture of that resident's capacity and needs for improvement. True clinical competence is the ability to use a complete set of skills dynamically, based on the needs of a specific situation.<sup>1,3</sup>

Simulation poses a unique opportunity to create an artificial environment that allows for broader assessment of resident skill.<sup>4</sup> The observed structured clinical examination (OSCE) can generate reliable and accurate measures of trainee ability, especially in the domains of communication and professionalism skills.<sup>10,11</sup> Trained standardized patient raters can identify deficiencies that experts miss because of fewer biases, and may

better appreciate the patient's perspective.<sup>12</sup> It has been argued that OSCEs are limited in their ability to assess higher levels of clinical competence.<sup>1</sup> OSCEs must be carefully designed and implemented, especially if the results are to be used for high-stakes decision-making, such as promotion in a training program.<sup>2,12,13</sup>

Educators of trainees in procedure-based specialties need an assessment tool that can be modified to address a variety of technical skills for learners at varying levels, within a more realistic environment, to assess the true capability of the physician-in-training. A variety of useful scales have been developed for measurement of specific skills, for instance, laparoscopic skills,<sup>9</sup> but unique scales for each type of procedure make composite measures difficult to obtain. There have been several efforts to create hybrid OSCEs that include technical skills testing within a case-based scenario that might better approximate reality, and require the trainees to call on a range of skills.<sup>15-17</sup> A hybrid examination that provided a framework into which a variety of skills could be assessed would be useful in training residents. In obstetrics and gynecology, residents perform procedures in the delivery room, as well as laparoscopic, hysteroscopic, and open surgery. We set out to create a simple framework that might allow for bench skills testing within the OSCE structure, and could be modified for a variety of different skills.

Repair of episiotomies and vaginal lacerations is a core skill that all obstetrician/gynecologists must master in their training. Generally introduced to junior residents early in their training, this is a common and important procedure. Of women who have vaginal delivery, 85% have perineal trauma spontaneously or because of an episiotomy, and the majority require suturing to facilitate healing.<sup>18-20</sup> Incorrectly repaired perineal trauma can result in persistent pain, anal incontinence and even lead to increased postpartum depression.<sup>21</sup> Episiotomy models have been developed and used for resident training and assessment.<sup>22</sup> We incorporated a vaginal laceration repair into a resident OSCE, videotaped the procedure, and used a modified global assessment scale to rate the skills. This is a description of our evaluation tool and early performance data.

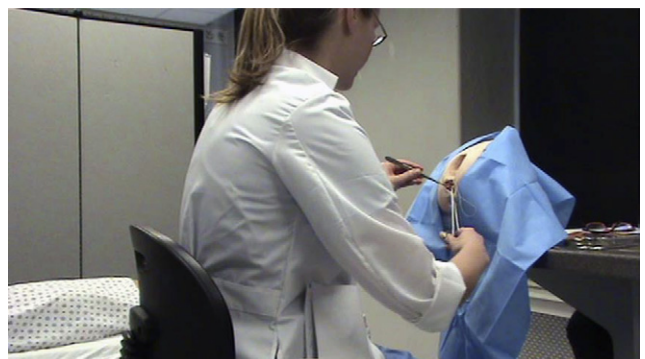
## METHODS

We created a 5-station hybrid OSCE with standardized patient scenarios and partial task manikins for the first-year residents in obstetrics and gynecology in the 11th month of their first year of residency. All the residents had participated in 6 months of obstetrics rotations by the time of the OSCE. The 5 stations included scenarios of a patient who had just delivered a baby and required a vaginal laceration repair—the subject of this study—and 4 others that included management of menorrhagia with a dilation and curettage, a family meeting of a patient with terminal ovarian cancer, a new diagnosis of a pregnancy loss, and a review of the medical literature to determine appropriate antibiotic therapy for pelvic inflammatory disease. Each of the stations of the OSCE involved actors trained as standardized patients as well as raters of the residents, as described in

previous reports from our institution.<sup>10,11</sup> This checklist included behaviorally anchored items addressing information gathering, relationship development, patient education and counseling, organization and time management, focused history-taking, cultural competency managing difficult situations, and patient satisfaction. Some of these items were assessed across each case in the OSCE, and some items were case-specific. A faculty member observed each station. Each station consisted of 15 minutes for the patient encounter, including the laceration repair. Faculty participated in debriefing and provided formative feedback for 5 minutes after every station. Because the 5 faculty members participating in the examination knew the residents and had not been formally trained as raters, their feedback was purely formative.

In the vaginal laceration scenario, the standardized patient presented after having just delivered a baby in an ambulance. She had delayed presenting to care because she was an undocumented immigrant, and called 9-1-1 when she was in active labor. The residents were informed by an instruction sheet on the door to discuss the situation with the patient and attend to the necessary repair of a second degree vaginal laceration on a model (Episiotomy Trainer 60,225; Limbs and Things, Savannah, GA). The model was selected after review by 2 attending obstetrician-gynecologists.

A stationary video camera was set up to face the model (Figure 1: still image of laceration repair), and the procedure repair was recorded for later evaluation by 2 attending obstetrics and gynecology faculty members invested in the assessment and trained as raters. Resident performance was rated using a modified version of the Global Rating Scale of Operative Performance, a 5-item behaviorally anchored scale for procedural assessment validated in a variety of surgical training milieu, in and out of the operating room.<sup>7,9,14</sup> (Figure 2: modified assessment scale) Residents were rated on a scale from 0 to 2, with 0 representing “not done,” 1 representing “partly done,” and 2 representing “well done;” this is consistent with the checklist scoring for the communication and professionalism domains of the OSCE.<sup>23,24</sup> Two attending obstetrician-gynecologists independently rated the skills, and the results were averaged, and analyzed using SPSS ver. 19, SPSS Inc., Chicago, IL. The residents completed questionnaires about the overall examination and specific stations. The results of the technical portion of the



**FIGURE 1.** Still image of laceration repair.



Obstetrics & Gynecology IPSA 2011  
Procedure evaluation form

Resident ID: \_\_\_\_\_

Attending Name: \_\_\_\_\_

Station: \_\_\_\_\_

	1	2	3
<b>Respect for tissue</b>	Used unnecessary force on tissue or caused damage by inappropriate use of instruments.	Careful handling of tissue, but occasionally caused inadvertent damage.	Consistently handled tissues appropriately with minimal damage.
<b>Time and motion</b>	Many unnecessary moves.	Efficient time/motion, but some unnecessary moves.	Economy of movement and maximum efficiency.
<b>Instrument handling</b>	Repeatedly makes tentative or awkward moves with instruments.	Competent use of instruments although occasionally appeared stiff or awkward.	Fluid moves with instruments and no awkwardness.
<b>Knowledge of instruments</b>	Frequently used an inappropriate instrument.	Usually used appropriate instruments for the task.	Obviously familiar with the instruments required.
<b>Knowledge of specific procedure and patient</b>	Deficient knowledge. Needed specific instruction at most operative steps.	Knew all important aspects of the operation, had some gaps in knowledge of case history or workup.	Demonstrated familiarity with all aspects of the operation and patient history.

Comments: \_\_\_\_\_

- Overall Assessment: Needs improvement  
Competent to perform with supervision  
Competent to perform independently

**FIGURE 2.** Modified Assessment Scale.

examination were compiled and shared with residents during their semiannual performance evaluations.

All residents in the study participated in the Research on Medical Education Outcomes (ROME) Registry, maintained at NYU and supported by the Institutional Review Board (IRB). The residents' identifying information is removed for the purpose of analysis. This study was submitted to the IRB and classified as exempt from IRB review because it was part of the educational program.

## RESULTS

### Resident Evaluation

Seven first-year residents in obstetrics and gynecology completed the assessment. Mean performance of the task elements

as "well-done" was 55% (range 20%-90%). There were 3 residents who performed below the mean, and their performance was further reviewed. Of these residents, 2 had scores of "partly done" on all the other elements, while only 1 resident had "not done" several required elements of the task (Table 1). Reliability analysis of the data was done, which yielded a Cronbach's alpha rating of 0.817 for the station. While 1 resident scored poorly across several elements of the hybrid examination, for the most part resident performance varied between technical performance, communication, and professionalism (Table 2).

### Program Evaluation

Subjective comments about the scenario from the residents endorsed finding it difficult to juggle bonding with the patient, reassuring her about the well-being of herself and her infant,

**TABLE 1.** Resident Performance of Laceration Repair

	<b>Skill Elements Well-Done</b>	<b>Deviation from Mean</b>
Resident 1	30%	-25%
Resident 2	70%	+15%
Resident 3	90%	+35
Resident 4	80%	+25%
Resident 5	20%	-35%
Resident 6	60%	+5%
Resident 7	40%	-15%

Based on scoring of videotaped performance by 2 faculty members. Mean performance was 55% "Well-Done" with a standard deviation of 26%. Residents with performance falling below the mean are highlighted in gray.

and attending to the bleeding laceration that needed repair (3 comments). One resident reported feeling uncomfortable with the difficult situation of the patient's uncertain immigration status. There were technical comments, such as the position of the bench model relative to the patient (2 comments) and the foam on the model being damaged by earlier repairs by the time the final resident rotated through the scenario.

## CONCLUSIONS

By combining the OSCE format with bench models for skills testing, we created a framework in which to assess the building blocks of competence together. While educators will continue to argue about whether performance in the simulated environment accurately represents clinical competence,<sup>1,13</sup> the pressure on educators to create benchmarks for learning and objective tools for assessment is not waning. Rather, it is becoming increasingly important to document competence in trainees. A framework for hybrid OSCEs that could assess a variety of technical skills would be useful for educators looking to assess resident ability across a range of competencies.

In this study, we created a scenario that called upon trainees' skills in rapport-building, cultural competency, and patient counseling as well as technical ability to repair a vaginal laceration. The simple design and evaluation tool could be modified to create a different scenario or assess different technical skills depending on the needs of the training program, or the level of

the learner. The combined feedback from all the stations of the examination yields different results for each resident across a range of competencies, which can help to establish benchmarks for residents at different levels of training.

Scores reflect what we would expect for first year residents: that the group performed the task appropriately with some deficiencies. Interestingly, residents who performed poorly on technical skills did not necessarily perform poorly on communication or professionalism measures. Residents performing well below the mean in a certain area were directed to those areas for focusing their own learning goals. Group performance was lower on the domain of professionalism, which is a useful needs assessment for the residency program, and might help guide curriculum development.

The scores of residents performing below the mean in technical ability were further reviewed, and revealed that 1 resident had poor performance in the domains of instrument handling and respect for tissue. This resident could be directed to focus learning efforts on these skills, as similar deficiencies were not noted on cognitive or affective elements of the assessment. This can be interpreted as evidence of the sensitivity of the assessment tool, although this needs to be further evaluated in subsequent installments of the assessment with larger numbers of learners. The subjective comments are enlightening, and help with further evolution of the examination. Residents performing below the mean were given the opportunity to review their performance on video with a faculty member. The comments from the residents about difficulty managing time between the technical repair and the patient's complex social and emotional situation reinforce the rationale behind this project. As in real life, the residents were called upon to balance cultural competency, patient education, and counseling regarding procedure with the technical procedure itself.

This is a small pilot study and, as a result, the numbers are not statistically significant. We could not compare this technical assessment with other measures of the residents' in vivo clinical ability to validate the measure. While attendings are expected to complete evaluation forms for every procedure done by residents in our program, this is rarely done. This is a limitation of our study, but also underscores the need to investigate means of assessment of residents that yields valid, objec-

**TABLE 2.** Resident Performance Across Competencies

	<b>Technical Skills (Mean 55%)</b>	<b>Communication (Mean 67%)</b>	<b>Professionalism (Mean 25%)</b>
Resident 1	-25%	-22%	-25%
Resident 2	+15%	+21%	+55%
Resident 3	+35	-11%	-25%
Resident 4	+25%	-3%	-25%
Resident 5	-35%	-3%	+15%
Resident 6	+5%	+2%	-5%
Resident 7	-15%	+15%	+15%

Deviation from mean score of group. Communication and Professionalism scores derived from behaviorally anchored checklists completed by standardized patients trained as raters. Technical skills scores derived from videotaped performance of technical skills rated by two faculty members. Scores greater than 5% above or below the mean are highlighted in gray.



tive, instructive feedback. It is worth noting that the episiotomy trainer was selected because of ease of use and storage, as well as the manufacturer's promise of multiple uses. However, after being used by 7 examinees, the foam and plastic elements had broken down. We find the beef tongue model sometimes used for laceration repair training to be more realistic and cheaper although more labor-intensive to assemble.

The goal of this work was to see whether skills testing could be incorporated into an OSCE framework, to create a testing environment, which brings more complex dimensions that come closer to approximating the actual practice of medicine and surgery. The overall scores help the residency program to assess its needs, and the individual results help each resident to individualize learning goals. The feedback from this scenario was included into a "report card" for the residents, which helped the program directors discuss each resident's strengths and weaknesses during the examination. There is potential for this model to be adapted to other skills, and refined and validated to help standards for competency.

## REFERENCES

1. Huddle TS, Heudebert GR. Viewpoint: taking apart the art: the risk of anatomizing clinical competence. *Acad Med.* 2007;82:536-541.
2. Kennedy TJ, Regehr G, Baker GR, Lingard LA. Progressive independence in clinical training: A tradition worth defending? *Acad Med.* 2005;80 (10 Suppl):S106-S111.
3. Wanzel KR, Ward M, Reznick RK. Teaching the surgical craft: From selection to certification. *Curr Probl Surg.* 2002;39:573-659.
4. Greco EF, Regehr G, Okrainec A. Identifying and classifying problem areas in laparoscopic skills acquisition: Can simulators help? *Acad Med.* 2010;85:S5-S8.
5. Cox M, Irby DM. Assessment in medical education. *N Engl J Med.* 2007;356:387-396.
6. Vassiliou MC, Feldman LS. Objective assessment, selection, and certification in surgery. *Surg Oncol.* 2011;20:140-145.
7. van Hove PD, Tuijthof GJ, Verdaasdonk EG, Stassen LP, Dankelman J. Objective assessment of technical surgical skills. *Br J Surg.* 2010;97:972-987.
8. Martin JA, Regehr G, Reznick R, et al. Objective structured assessment of technical skill (OSATS) for surgical residents. *Br J Surg.* 1997;84:273-278.
9. Vassiliou MC, Feldman LS, Andrew CG, et al. A global assessment tool for evaluation of intraoperative laparoscopic skills. *Am J Surg.* 2005;190:107-113.
10. Hochberg MS, Kalet A, Zabar S, et al. Can professionalism be taught? Encouraging evidence. *Am J Surg.* 2010;199:86-93.
11. Zabar S, Hanley K, Stevens DL, et al. Can interactive skills-based seminars with standardized patients enhance clinicians' prevention skills? Measuring the impact of a CME program. *Patient Educ Couns.* 2010;80:248-252.
12. Turner JL, Dankoski ME. Objective structured clinical exams: A critical review. *Fam Med.* 2008;40:574-578.
13. Ginsburg S, McIlroy J, Oulanova O, et al. Toward authentic clinical evaluation: Pitfalls in the pursuit of competency. *Acad Med.* 2010;85:780-786.
14. Dath D, Regehr G, Birch D, et al. Toward reliable operative assessment: The reliability and feasibility of videotaped assessment of laparoscopic technical skills. *Surg Endosc.* 2004;18:1800-1804.
15. Maker VK, Bonne S. Novel hybrid objective structured assessment of technical skills/objective structured clinical examinations in comprehensive perioperative breast care: A three-year analysis of outcomes. *J Surg Educ.* 2009;66:344-351.
16. Alinier G. Nursing students' and lecturers' perspectives of objective structured clinical examination incorporating simulation. *Nurse Educ Today.* 2003;23:419-426.
17. Ruesseler M, Weinlich M, Byhahn C, et al. Increased authenticity in practical assessment using emergency case OSCE stations. *Adv Health Sci Educ Theory Pract.* 2010;15:81-95.
18. McCandlish R, Bowler U, van Asten H, et al. A randomized controlled trial of care of the perineum during second stage of normal labor. *Br J Obstet Gynaecol.* 1998;105:1262-1272.
19. da Silva FM, de Oliveira SM, Bick D, et al. Risk factors for birth-related perineal trauma: A cross-sectional study in a birth centre. *J Clin Nurs.* 2012.
20. Aasheim V, Nilsen AB, Lukasse M, Reinart LM. Perineal techniques during the second stage of labor for reducing perineal trauma. *Cochrane Database Syst Rev.* 2011;12:CD006672.
21. Bahl R, Murphy DJ, Strachan. Non-technical skills for obstetricians conducting forceps and vacuum deliveries: Qualitative analysis by interviews and video recordings. *Eur J Obstet Gynecol Reprod Biol.* 2010;150:147-151.
22. Banks E, Pardanani S, King M, et al. A surgical skills laboratory improves residents' knowledge and performance of episiotomy repair. *Am J Obstet Gynecol.* 2006;195:1463-1467.
23. Chander B, Kule R, Baiocco P, et al. Teaching the competencies: Using objective structured clinical encounters for gastroenterology fellows. *Clin Gastroenterol Hepatol.* 2009;7:509-514.
24. Regehr G, MacRae H, Reznick RK, Szalay D. Comparing the psychometric properties of checklists and global rating scales for assessing performance on an OSCE-format examination. *Acad Med.* 1998;73:993-997.