

## EDUCATION

**Objective assessment of vaginal surgical skills**

Chi Chiung Grace Chen, MD; Abner Korn, MD; Christopher Klingele, MD; Matthew D. Barber, MD, MHS; Marie Fidela R. Paraiso, MD; Mark D. Walters, MD; J. Eric Jelovsek, MD

**OBJECTIVE:** To develop and validate an instrument to assess surgical skills during vaginal surgery.

**STUDY DESIGN:** Trainees from 2 institutions were directly assessed in the operating room by supervising surgeons while performing a vaginal hysterectomy using the new Vaginal Surgical Skills Index, global rating scale, and visual analogue scale. Trainees were assessed again by the same surgeons 4 weeks after the live surgery and by a blinded outside reviewer using a videotape of the case. Internal consistency, interrater and intrarater reliability, and construct validity were evaluated.

**RESULTS:** Two hundred twelve evaluations were analyzed on 76 surgeries from 27 trainees. There was good internal consistency, interrater,

and intrarater reliability. Vaginal Surgical Skills Index scores correlated with global rating score and visual analog scale scores. Increasing Vaginal Surgical Skills Index scores significantly correlated with year of training and surgical volume with an estimated increase in score of 0.3 per hysterectomy performed.

**CONCLUSION:** The Vaginal Surgical Skills Index is a feasible, reliable, and valid instrument to assess vaginal surgical skills.

**Key words:** global rating scale, reliability, surgical skills evaluation, vaginal surgery, validity

Cite this article as: Chen C, Korn A, Klingele C, et al. Objective assessment of vaginal surgical skills. *Am J Obstet Gynecol* 2010;203:79.e1-8.

Despite advances in surgical instruments and techniques, progress in formalized assessment of surgical skills for obstetrics and gynecology residents is limited. A 1997-1998 survey of obstetrics and gynecology residency directors in the United States found that only 29% of

programs had any formal surgical curricula and only 17% used any method of standardized assessment of resident technical skills.<sup>1</sup> Since this survey, many residency programs have improved their assessment strategies; however, assessments of surgical skills continue to be limited to nonvalidated global assessments by supervising surgeons at the end of a surgical rotation and by the use of surgical case logs. Case logs lack content validity because they only document participation in a case and do not assess the technical ability of the individual performing the procedure and most forms of global assessment have poor reliability and unknown validity.<sup>2</sup> Although surgical checklists are also useful for learning and giving feedback, for assessment purposes, global rating scales scored by experts have higher reliability and validity than checklists. However, there are currently no published assessment instruments that have been tested for reliability or validity for use in live open gynecologic or vaginal surgery.

Formal intraoperative assessment is difficult to standardize because the difficulty level of the procedure and the amount of participation from each trainee during the case is variable. Cadaveric specimens, live animals, and bench top simulators have been used in an attempt to standardize the

assessment of basic surgical skills outside the operating room. Objective Structured Assessments of Technical Skills in these settings have good reliability and validity using animal models and bench top simulators,<sup>3-5</sup> yet they are limited by their availability, cost, and need for faculty time. Moreover, facilities required to support these programs may not be feasible for smaller community training programs. Therefore, the gold standard of determining a surgeon's competence remains assessment during live surgery and valid and reliable intraoperative assessment tools should be developed and validated in gynecology for this purpose. These tools could also be beneficial in providing accurate documentation of a trainee's performance through time allowing for directed teaching and could be used to provide a template for providing constructive formative feedback about their performance.

Intraoperative assessment instruments should be psychometrically robust, feasible to administer, and easily adaptable by residency programs. The most commonly used instrument to evaluate surgical skills is the global rating of operative performance developed by Dr Reznick and his general surgery colleagues.<sup>6,7</sup> This instrument consists of a 7-item global rating scale that allows supervising surgeons to rate important but

From the Obstetrics, Gynecology and Women's Health Institute, Cleveland Clinic (Drs Chen, Barber, Paraiso, Walters, and Jelovsek), Cleveland, OH; the Department of Obstetrics and Gynecology (Dr Korn), University of California, San Francisco, School of Medicine, San Francisco, CA; and the Department of Obstetrics and Gynecology (Dr Klingele), Mayo Clinic, Rochester, MN.

Presented at the 35th Annual Scientific Meeting of the Society of Gynecologic Surgeons, New Orleans, LA, March 30-April 1, 2009.

Received Aug. 13, 2009; revised Sept. 1, 2009; accepted Jan 24, 2010.

Reprints: J. Eric Jelovsek, MD, Cleveland Clinic, Department of Obstetrics and Gynecology, Building A81, 9500 Euclid Ave., Cleveland, OH 44195. [jelovsj@ccf.org](mailto:jelovsj@ccf.org).

Financial support was provided by the Multi-Center Education Research Award, Society of Gynecologic Surgeons.

0002-9378/\$36.00

© 2010 Mosby, Inc. All rights reserved.

doi: 10.1016/j.ajog.2010.01.072

generic skills during surgical procedures including: respect for tissue, time and motion, instrument handling, knowledge of instruments, use of assistants, flow of operation, and knowledge of the procedure. Although this performance index has been shown to be reliable and valid when used with bench top simulators and animal models, its psychometric properties have not been demonstrated using blinded reviewers in live surgery, despite it being the only tool listed on the Accreditation Council for Graduate Medical Education Patient Care Assessment section for the assessment of surgical skills. Furthermore, although, this instrument has been used in the gynecologic literature, albeit sparingly, its psychometric properties have not been determined for live surgery and there are currently no reliable and valid instruments for assessing vaginal surgical skills in the laboratory or in the operating room setting.

The primary aims of this study were to develop a feasible, reliable, and valid instrument to evaluate surgical skills during vaginal surgery and to compare the psychometric properties of this new assessment scale for vaginal surgery with the more commonly used Global Rating Scale (GRS) of Operative Performance. Our goal was to develop an assessment tool that is representative of essential skills required to perform vaginal surgery. We hypothesized that the Vaginal Surgical Skills Index (VSSI) would be a feasible, reliable, and valid instrument for the assessment of vaginal surgical skills while performing vaginal hysterectomy.

## MATERIALS AND METHODS

To develop an instrument specifically for the evaluation of vaginal surgical skills, we expanded and modified the original 7-item GRS. Additional items were included such as initial inspection, performance of an incision, maintenance of visibility, use of electrosurgery, knot tying, maintenance of hemostasis, removing fluid and debris at the completion of the procedure, and forward planning for a total of 13 items (Appendix). These additional items were deemed important and highly relevant during vaginal surgery

TABLE 1

### Internal consistency of the VSSI and the GRS of operative performance during vaginal surgery by method of evaluation

Evaluation method	VSSI <sup>a</sup>	GRS <sup>a</sup>
Live surgery	0.96	0.95
Blinded videotape review by the same surgeon after 1 mo	0.97	0.94
Blinded external videotape review	0.95	0.94

GRS, Global Rating Scale; VSSI, Vaginal Surgical Skills Index.

<sup>a</sup> Cronbach's alpha.

Chen. Objective assessment of vaginal surgical skills. *Am J Obstet Gynecol* 2010.

based on feedback from practicing gynecologic surgeons given the unique aspects of operating using a vaginal approach. The descriptors of the GRS were also rewritten to provide clear descriptions for the assessor of what was to be rated and these descriptors were added to all items rather than every other item on the GRS. To rate the trainee's overall performance during the procedure, a 100-mm visual analog scale (VAS) was also included as an additional measure of the attending's global impression of surgical skill by the trainee. Content and face validity of the VSSI were assessed by surveying gynecology trainees and a panel of teaching gynecologic surgeons from 3 training institutions as to whether the scale appeared fair and appropriate in assessing important qualities of surgical technique during vaginal surgery. Because this project was presented to, discussed at, and funded by the Society of Gynecologic Surgeons Education Research Forum in 2006, we also benefited from input from a wider group of gynecologic surgeons from multiple training institutions.

Institutional Review Board exemption was obtained as this project involved the use of educational tests and did not affect the clinical course of the patients. We eval-

uated obstetrics and gynecology residents and postresidency, Female Pelvic Medicine and Reconstructive Surgery fellows during the performance of a vaginal hysterectomy at the Cleveland Clinic and Mayo Clinic from May 2007 to June 2008 using the new VSSI, GRS, and VAS. Trainees were evaluated several times over the course of their gynecology rotation. Trainees were instructed to perform the hysterectomy as if they were the primary surgeon and neither the assistant nor the supervising surgeon would interfere with decisions as long as the patient's safety was maintained. The vaginal hysterectomy was videotaped from the incision to completion with removal of the uterus and inspection of all pedicles for hemostasis while preserving the anonymity of the trainee. As the vaginal hysterectomy was often performed with other procedures such as salpingo-oophorectomy or prolapse surgeries, we did not include vaginal cuff closure in the video recording. Because the video recording did not include audio, the supervising attending and other surgical assistants were instructed not to assist or correct the trainee verbally; instead, every effort was made to correct or guide the trainee with hand gestures so that it was clear on the

TABLE 2

### Interrater reliability and intrarater reliability for the VSSI, GRS of operative performance, and VAS during vaginal surgery

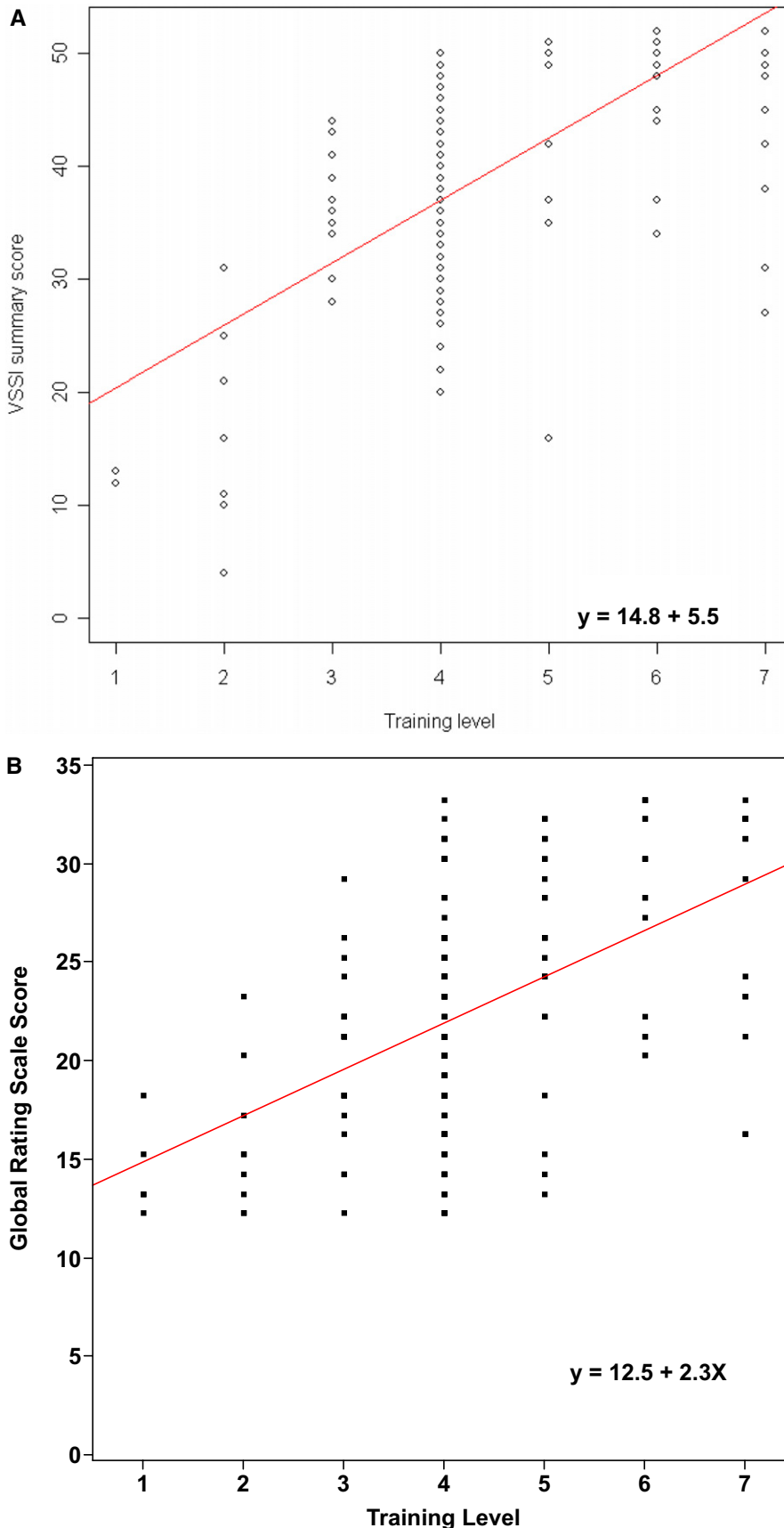
Assessment scale	Interrater reliability <sup>a</sup>	Intrarater reliability <sup>a</sup>
VSSI	0.53	0.82
GRS	0.31	0.64
VAS	0.51	0.69

GRS, Global Rating Scale; VAS, visual analog scale; VSSI, Vaginal Surgical Skills Index.

<sup>a</sup> Intraclass correlation coefficient.

Chen. Objective assessment of vaginal surgical skills. *Am J Obstet Gynecol* 2010.

**FIGURE 1**  
Univariate relationship between live summary scores and training level



video recording that the trainee had made an error and was being rectified or the trainee did not know the steps of the procedure. Assistants were also instructed not to automatically assist but to have the trainee demonstrate what assistance was needed because the use of assistance was 1 of the evaluation parameters on the VSSI form. The VSSI, GRS, and VAS were completed by a supervising attending surgeon immediately after observing the trainee perform a vaginal hysterectomy. Additional information collected included the time taken to perform the hysterectomy and to fill out the VSSI form, training level of the trainee, and number of vaginal hysterectomies the trainee had previously performed. The attending surgeon also rated the difficulty of the case by answering the question, “Compared with average level of difficulty for your operations, would you say this operation was?” and checking a 5-point discrete response scale ranging from “much easier than my average operation” to “much more difficult.”

To evaluate interrater reliability, a copy of the blinded video-recorded procedure was sent to a surgeon at a separate institution who evaluated the trainee’s performance using the same assessment forms after watching the procedure. Another copy of the recorded procedure was given to the intraoperative supervising physician for review and form completion at least 4 weeks after the date of the original procedure to assess intrarater reliability. Both evaluations were performed in a blinded manner with the anonymity of the trainee and their training level preserved as the recordings only included the date of surgery and the trainee 4-digit unique identifier.

Cronbach’s alpha was calculated to assess internal consistency of the VSSI, assuming independence across surgeries. In-

**A**, Univariate relationship between the VSSI summary scores and training level indicating good discriminate validity. **B**, Univariate relationship between the GRS of operative performance summary scores and training level indicating good discriminate validity.

GRS, Global Rating Scale; VSSI, Vaginal Surgical Skills Index.

Chen. Objective assessment of vaginal surgical skills. *Am J Obstet Gynecol* 2010.

**TABLE 3**  
**Convergent validity between the VSSI, GRS,**  
**and VAS by method of evaluation**

Evaluation method	VSSI and VAS <sup>a</sup>	VSSI and GRS <sup>a</sup>
Live surgery	0.91	0.94
Blinded videotape review by the same surgeon after 1 mo	0.93	0.92
Blinded external videotape review	0.92	0.94

GRS, Global Rating Scale; VAS, visual analog scale; VSSI, Vaginal Surgical Skills Index.

<sup>a</sup> Pearson correlation coefficient.

Chen. Objective assessment of vaginal surgical skills. *Am J Obstet Gynecol* 2010.

terater reliability and intrarater reliability were analyzed using Intraclass correlation coefficient (ICC). Construct validity was evaluated by measuring convergent validity, or the association between VSSI scores (range, 0–52) and the GRS scores (range, 0–35) or VAS scores (range, 0–100) using Pearson correlation coefficient ( $r$ ) and by discriminant validity by comparing VSSI scores with training level and surgical volume. Because most trainees had multiple surgeries assessed, the estimated mean and standard error of the instrument scores by training level were compared. Based on previous studies validating the GRS, at least 20 trainees would be necessary to discriminate between training levels to demonstrate discriminant validity.<sup>6</sup> Linear mixed effect models were used to evaluate the relationship between the VSSI scores and training level or surgical volume while considering the random trainee effect and the correlation within surgery. All analyses were performed using SAS 9.1 software (SAS Institute, Inc, Cary, NC).

## Results

A total of 27 trainees postgraduate training levels (PGY) 1–7 (residents and fellows) from 2 institutions participated in this study. Because most trainees had multiple procedures assessed during the study period, there were a total of 76 surgeries performed. Five supervising surgeons from 2 institutions (C.C. and M.C.) completed evaluations immediately after each procedure and then again after watching the recorded procedure. One surgeon from a third institution (University of California, San Francisco) served as the blinded external reviewer who watched all the recorded procedures and completed the evaluations for each

of the procedures. A total of 212 surgical evaluations were analyzed. The mean ( $\pm$  standard deviation [SD]) time taken to complete vaginal hysterectomy was 36.6  $\pm$  17.8 minutes (range, 8–135) and the mean time taken to complete VSSI form was 2  $\pm$  1 minutes (range, 1–10).

Internal consistency for the VSSI and GRS were high (Cronbach's alpha = 0.95–0.97). As shown in Table 1, interrater reliability (ICC = 0.53) and intrarater reliability (ICC = 0.82) for the VSSI were also good; however, the GRS had the lowest interrater reliability at 0.31 and intrarater reliability was 0.64 (Table 2).

Both scales were able to discriminate between training levels as there was a moderate correlation between the VSSI total score and training level  $r = 0.64$  (95% confidence interval [CI], 0.53–0.73;  $P < .001$ ) (Figure 1, A) and between the GRS scores and training level ( $r = 0.51$ ; 95% CI, 0.40–0.61;  $P < .001$ ) (Figure 1, B). The VSSI and GRS also appeared to have good convergent validity as the correlations between the overall VSSI score, VAS, and overall GRS score were high, ranging from 0.90–0.94 (Table 3). After controlling for differences in method of rating and level of surgery difficulty, increasing VSSI scores and GRS continued to significantly correlate with training level with each year corresponding to an average increase of 6.6 (95% CI, 4.2–7.0;  $P < .001$ ) points on the total VSSI score per year of training and an average increase of 2.54 (95% CI, 1.96–3.23) points on the total GRS score for each additional surgery performed. Surgical volume also significantly correlated with live VSSI scores (Figure 2, A) and live GRS scores (Figure 2, B). After controlling for

differences in method of rating and level of surgery difficulty, there was an estimated increase of 0.3 points in the VSSI score (parameter estimate, 0.27; 95% CI, 0.20–0.33;  $P < .001$ ) for each additional vaginal hysterectomy performed and 0.13 in the mean score of the GRS for each additional hysterectomy performed (parameter estimate, 0.13; 95% CI, 0.10–0.16;  $P < .001$ ).

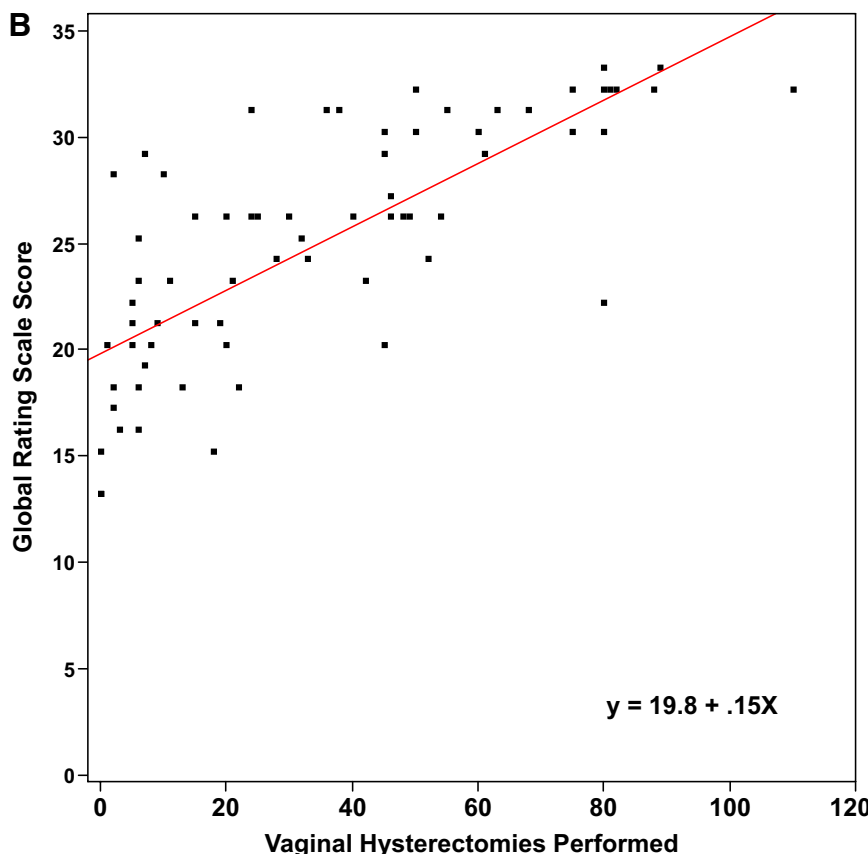
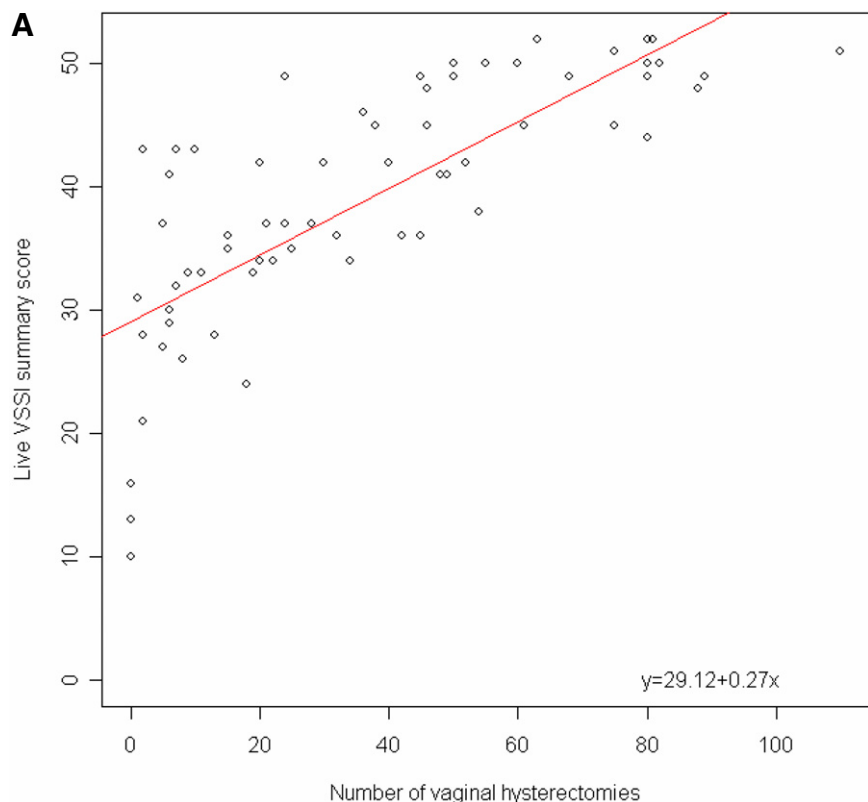
## Comment

There has been growing concern among obstetrics and gynecology educators that with the increased requirements for primary care and the restrictions on work hours, surgical learning may be compromised.<sup>8,9</sup> Accrediting bodies continue to implement requirements that training programs provide valid and reliable documentation of surgical performance of their trainees and accurately attest to an individual's surgical abilities before completion of training. Historically, programs have relied on subjective evaluations of a resident's skills. Assessments of this nature are often biased, inconsistent, unreliable, and invalid.<sup>2,10</sup> Because of these limitations, these evaluations arguably add little to a trainee's understanding of their abilities and areas that need improvement. It is therefore timely and imperative to establish more valid and reliable assessment strategies to evaluate surgical techniques.

There are 2 major differences between using an assessment tool in a controlled environment like the Objective Structured Assessments of Technical Skills compared with live surgery. In live surgery, assessors cannot be blinded to the trainee, patient characteristics, or operating room environment. Knowledge of any of these factors may be a source of bias into the assessment. Second, live cases cannot be completely standardized as the difficulty level of the case will be variable. Difficulty level introduces variability into the trainee's performance. Therefore, for an assessment tool to be an appropriate measure of a live procedure it must be accurate enough to overcome this variability to be a reliable measure.

Our goal was to develop a valid and reliable instrument to measure performance of vaginal surgery given its estab-

**FIGURE 2**  
**Univariate relationship between live summary scores and number of vaginal hysterectomies**



lished efficacy and safety in gynecologic surgery. In 2003, more than 600,000 hysterectomies were performed in the United States with 22% of these performed using the vaginal approach.<sup>11</sup> Vaginal surgery is considered the approach of choice for most patients requiring hysterectomy because morbidity is lower with the vaginal approach than with any other method.<sup>12,13</sup> Furthermore, vaginal surgery requires a unique set of surgical skills because of limited exposure, limited assistance, and operating in a narrow cavity. Therefore, it is essential to have a specific instrument for assessment of vaginal surgical skills.

In this study, we successfully demonstrated good interrater and intrarater reliability and construct validity of the VSSI. The GRS demonstrated high internal consistency and moderately high intrarater reliability. The interrater reliability for the GRS is low in this study and also lower than the VSSI. This is concerning and may be explained by the blinded study design and assessment of live, rather than standardized surgery. Most data describing the reliability of the GRS do not use blinded comparisons of live surgery; instead these use standardized stations of different surgical techniques in a surgical laboratory. By using live surgery, we included the effect of different severities of disease and patient variability on the performance. This potentially increases the variability of an assessor rating performance. However, this did not appear to affect the VSSI as the interrater reliability remained acceptable, despite the increased variability of rating live surgery.

Strengths include the prospective nature of the study and the attention to relevant psychometric properties for assessment instruments. The VSSI appears to be a valid and reliable assessment instrument for evaluating trainees per-

**A**, Univariate relationship between live VSSI summary scores and number of vaginal hysterectomies. **B**, Univariate relationship between live GRS of operative performance summary scores and number of vaginal hysterectomies.

GRS, Global Rating Scale; VSSI, Vaginal Surgical Skills Index.

Chen. Objective assessment of vaginal surgical skills. *Am J Obstet Gynecol* 2010.

forming vaginal hysterectomy. One of the most important properties of any assessment scale is its reliability (ie, the tool should measure the same thing each time it is used under the same conditions with the same subjects.) We have shown that the VSSI demonstrated high internal consistency and moderately high interrater reliability. These are reassuring properties that allow for trainees to be compared across different evaluators or across institutions; potentially using video-taped assessments. The VSSI also demonstrated good intrarater reliability indicating that the same reviewer would score the same trainee doing the same surgery the same way on different occasions. This is an important property to compare the progress of an individual over time by the same evaluating surgeons at an institution.

Construct validity was assessed using convergent and discriminate validity. Convergent validity is confirmed by agreement between different rating systems. For example, a good score on 1 rating system should also translate to a good score on another rating system if they are both measuring the same trainee doing the same series of tasks. In our study, VSSI scores highly correlated with GRS and VAS scores. The better a trainee did on the VSSI, the better they did on the GRS and VAS. Discriminate validity is the ability to distinguish between items that should be different. For example, a more senior trainee should have a different score than a more junior trainee. After controlling for differences in evaluation methods and difficulty of the surgery, there were significant differences in the VSSI scores among trainees of different levels with less experienced trainees receiving lower scores. Because trainees of the same level of training may have different surgical experiences, it is also important to understand the relationship between surgical volume and assessment scores. Therefore, in addition to discerning between different training levels, we further demonstrated construct validity by establishing a significant association between surgical volume and VSSI scores; higher surgical volume correlated with higher VSSI scores.

Assessment instruments should also be easy to fill out and considerate of a

surgeon's time. The VSSI is intended to be used after a trainee performs vaginal surgery and requires no special equipment or evaluator training. Most supervising surgeons were able to complete the form in 2 minutes, allowing the VSSI to be implemented and completed without significant impediment to the surgeon's time. Therefore, the VSSI could be easily incorporated into a residency training program for the evaluation and improvement of resident and fellow vaginal surgical skills.

One of the most surprising findings of the study was how much the process of assessment improved teaching during the study period. Most trainees gave positive remarks about being allowed to operate without unsolicited guidance during the cases. This was particularly emphasized in the PGY 3 and 4 residents. All trainees were eager to see how they performed on the VSSI and sought the immediate reinforcing or corrective feedback after the case. Using the VSSI as a guide to areas of surgical competency, the supervising surgeons explained to the trainee why they rated them the way they did and let them know specifically what they needed to change or learn more about. Although we did not specifically measure how much this affected learning, it is clear from this study that the VSSI can be not only used as a method of assessment but also as a teaching tool to give feedback and is the subject of a future study. Furthermore, this instrument should be tested over a trainee's full training of 4 years and at different institutions to assess its generalizability. It would also be important to establish competency cutoff scores on each instrument so that training programs could use to document surgical competency.<sup>14</sup>

In conclusion, the VSSI is a feasible, valid, and reliable assessment tool for vaginal surgery. This VSSI can be used to objectively evaluate a trainee's vaginal surgical skills to aid in the process of determining competency in vaginal surgery and it may also be used as a guide to give feedback immediately after surgical teaching to improve learning. The GRS is currently the only Accreditation Council for Graduate Medical Education recommended assessment tool for operative performance and these data suggest that

the instrument is a valid but less reliable measure of technical skills during the performance of vaginal surgery. A scale specific for vaginal surgery may be necessary to evaluate live surgery. ■

## REFERENCES

- Mandel LP, Lentz GM, Goff BA. Teaching and evaluating surgical skills. *Obstet Gynecol* 2000;95:783-5.
- Reznick RK. Teaching and testing technical skills. *Am J Surg* 1993;165:358-61.
- 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.
- Lentz GM, Mandel LS, Lee D, Gardella C, Melville J, Goff BA. Testing surgical skills of obstetric and gynecologic residents in a bench laboratory setting: validity and reliability. *Am J Obstet Gynecol* 2001;184:1462-8; discussion 1468-70.
- Goff BA, Lentz GM, Lee D, Fenner D, Morris J, Mandel LS. Development of a bench station objective structured assessment of technical skills. *Obstet Gynecol* 2001;98:412-6.
- Martin JA, Regehr G, Reznick R, et al. Objective structured assessment of technical skill (OSATS) for surgical residents. *Br J Surg* 1997;84:273-8.
- Reznick R, Regehr G, MacRae H, Martin J, McCulloch W. Testing technical skill via an innovative "bench station" examination. *Am J Surg* 1997;173:226-30.
- Sorosky JI, Anderson B. Surgical experiences and training of residents: perspective of experienced gynecologic oncologists. *Gynecol Oncol* 1999;75:222-3.
- Defoe DM, Power ML, Holzman GB, Carpentieri A, Schulkin J. Long hours and little sleep: work schedules of residents in obstetrics and gynecology. *Obstet Gynecol* 2001;97:1015-8.
- Watts J, Feldman WB, Neufeld VR, Norman GR, eds. *Assessing clinical competence*. New York: Springer; 1985: 259-74.
- Wu JM, Wechter ME, Geller EJ, Nguyen TV, Visco AG. Hysterectomy rates in the United States, 2003. *Obstet Gynecol* 2007;110:1091-5.
- American College of Obstetricians and Gynecologists. Appropriate use of laparoscopically assisted vaginal hysterectomy. ACOG committee on gynecologic practice bulletin no. 311. *Obstet Gynecol* 2005;105:929-30.
- Johnson N, Barlow D, Lethaby A, Tavender E, Curr E, Garry R. Surgical approach to hysterectomy for benign gynaecologic disease. *Cochrane Database Syst Rev* 2009;3:CD003677.
- Jelovsek JE, Waters MD, Korn A, et al. Establishing cutoff scores on assessments of surgical skills to determine surgical competence. *Am J Obstet Gynecol* 2010;203:81.e1-6.

## APPENDIX

## Vaginal Surgical Skills Index evaluation form

Trainee's name:

Procedure(s) the trainee performed:

1. \_\_\_\_\_
2. \_\_\_\_\_
3. \_\_\_\_\_

Date of operation:

Proportion of the procedure(s) that the trainee performed.

- ≤50% or >50%
- ≤50% or >50%
- ≤50% or >50%

Please evaluate each trainee according to the criteria below and check the box that most corresponds to their performance.

<b>1. Initial inspection</b> (Check one)	<input type="checkbox"/> 0 Incomplete and unsystematic inspection of relevant pelvic and vaginal structures	<input type="checkbox"/> 1 Partially complete and unsystematic inspection of relevant pelvic and vaginal structures	<input type="checkbox"/> 2 Complete but unsystematic inspection of relevant pelvic and vaginal structures	<input type="checkbox"/> 3 Complete and somewhat systematic inspection of relevant pelvic and vaginal structures	<input type="checkbox"/> 4 Systematic and complete assessment of relevant pelvic and vaginal structures	Not observed
<b>2. Incision</b> (Check one)	<input type="checkbox"/> 0 Does not perform appropriate incision(s) safely and does not use incision(s) effectively ensuring optimal exposure	<input type="checkbox"/> 1 Incompletely performs appropriate incision(s) safely and does not use incision(s) effectively ensuring optimal exposure	<input type="checkbox"/> 2 Performs appropriate incision(s) safely but does not use incision(s) effectively ensuring optimal exposure	<input type="checkbox"/> 3 Performs appropriate incision(s) safely and partially uses incision(s) effectively ensuring optimal exposure	<input type="checkbox"/> 4 Performs appropriate incision(s) safely and uses incision(s) effectively ensuring optimal exposure	Not observed
<b>3. Maintenance of visibility</b> (Check one)	<input type="checkbox"/> 0 Almost never or never obtains appropriate exposure	<input type="checkbox"/> 1 A few times (less than half the time) obtains appropriate exposure	<input type="checkbox"/> 2 Sometimes (about half the time) obtains appropriate exposure	<input type="checkbox"/> 3 Most time (more than half the time) obtains appropriate exposure	<input type="checkbox"/> 4 Almost always or always obtains appropriate exposure	Not observed
<b>4. Use of assistants</b> (Check one)	<input type="checkbox"/> 0 Almost never or never strategically used assistant(s) to the best advantage	<input type="checkbox"/> 1 A few times (less than half the time) strategically uses assistant(s) to the best advantage	<input type="checkbox"/> 2 Sometimes (about half the time) strategically uses assistant(s) to the best advantage	<input type="checkbox"/> 3 Most time (more than half the time) strategically uses assistant(s) to the best advantage	<input type="checkbox"/> 4 Almost always or always strategically uses assistant(s) to the best advantage at all times	Not observed
<b>5. Knowledge of instruments</b> (Check one)	<input type="checkbox"/> 0 Almost never or never uses and is familiar with correct instruments	<input type="checkbox"/> 1 A few times (less than half the time) uses and is familiar with correct instruments	<input type="checkbox"/> 2 Sometimes (about half the time) uses and is familiar with correct instruments	<input type="checkbox"/> 3 Most time (more than half the time) uses and is familiar with correct instruments	<input type="checkbox"/> 4 Almost always or always uses and is familiar with correct instruments	Not observed
<b>6. Tissue and instrument handling</b> (Check one)	<input type="checkbox"/> 0 Almost never or never appropriately handles tissue and instruments	<input type="checkbox"/> 1 A few times (less than half the time) appropriately handles tissue and instruments	<input type="checkbox"/> 2 Sometimes (about half the time) handles tissue and instruments appropriately	<input type="checkbox"/> 3 Most time (more than half the time) handles tissue and instruments appropriately	<input type="checkbox"/> 4 Almost always or always handles tissue and instruments appropriately	Not observed
<b>7. Electro-surgery</b> (Check one)	<input type="checkbox"/> 0 Almost never or never uses electro-surgery safely and efficiently	<input type="checkbox"/> 1 A few times (less than half the time) uses electro-surgery safely and efficiently	<input type="checkbox"/> 2 Sometimes (about half the time) uses electro-surgery safely and efficiently	<input type="checkbox"/> 3 Most time (more than half the time) uses electro-surgery safely and efficiently	<input type="checkbox"/> 4 Almost always or always uses electro-surgery safely and efficiently	Not observed

Chen. Objective assessment of vaginal surgical skills. Am J Obstet Gynecol 2010.

(continued)

## APPENDIX

## Vaginal Surgical Skills Index evaluation form (continued)

<b>8. Knot tying/ ligation (Check one)</b>	<input type="checkbox"/> 0 Almost never or never quickly and correctly performs suture ligation and knot tying	<input type="checkbox"/> 1 A few times (less than half the time) quickly and correctly performs suture ligation and knot tying	<input type="checkbox"/> 2 Sometimes (about half the time) quickly and correctly performs suture ligation and knot tying	<input type="checkbox"/> 3 Most time (more than half the time) quickly and correctly performs suture ligation and knot tying	<input type="checkbox"/> 4 Almost always or always quickly and correctly performs suture ligation and knot tying	Not observed
<b>9. Hemostasis (Check one)</b>	<input type="checkbox"/> 0 Almost never or never exposes bleeders and uses correct technique to obtain hemostasis safely and effectively	<input type="checkbox"/> 1 A few times (less than half the time) exposes bleeders and uses correct technique to obtain hemostasis safely and effectively	<input type="checkbox"/> 2 Sometimes (about half the time) exposes bleeders and uses correct technique to obtain hemostasis safely and effectively	<input type="checkbox"/> 3 Most time (more than half the time) exposes bleeders and uses correct technique to obtain hemostasis safely and effectively	<input type="checkbox"/> 4 Almost always or always exposes bleeders and uses correct technique to obtain hemostasis safely and effectively	Not observed
<b>10. Procedure completion (Check one)</b>	<input type="checkbox"/> 0 Almost never or never completely removes fluid and debris and thoroughly inspects for bleeding	<input type="checkbox"/> 1 A few times (less than half the time) completely removes fluid and debris and thoroughly inspects for bleeding	<input type="checkbox"/> 2 Sometimes (about half the time) completely removes fluid and debris and thoroughly inspects for bleeding	<input type="checkbox"/> 3 Most time (more than half the time) completely removes fluid and debris and thoroughly inspects for bleeding	<input type="checkbox"/> 4 Almost always or always completely removes fluid and debris and thoroughly inspects for bleeding	Not observed
<b>11. Time and motion (Check one)</b>	<input type="checkbox"/> 0 Almost never or never efficiently performs movements with no awkward or unnecessary moves	<input type="checkbox"/> 1 A few times (less than half the time) efficiently performs movements with no awkward or unnecessary moves	<input type="checkbox"/> 2 Sometimes (about half the time) efficiently performs movements with no awkward or unnecessary moves	<input type="checkbox"/> 3 Most time (more than half the time) efficiently performs movements with no awkward or unnecessary moves	<input type="checkbox"/> 4 Almost always or always efficiently performs movements with no awkward or unnecessary moves	Not observed
<b>12. Flow of operation and forward planning (Check one)</b>	<input type="checkbox"/> 0 Almost never or never demonstrates forward planning allowing for proper flow of the procedure	<input type="checkbox"/> 1 A few times (less than half the time) demonstrates forward planning allowing for proper flow of the procedure	<input type="checkbox"/> 2 Sometimes (about half the time) demonstrates forward planning allowing for proper flow of the procedure	<input type="checkbox"/> 3 Most time (more than half the time) demonstrates forward planning allowing for proper flow of the procedure	<input type="checkbox"/> 4 Almost always or always demonstrates forward planning allowing for proper flow of the procedure	Not observed
<b>13. Knowledge of specific procedure (Check one)</b>	<input type="checkbox"/> 0 Almost never or never demonstrates familiarity with all aspects of the operation	<input type="checkbox"/> 1 A few times (less than half the time) demonstrated familiarity with all aspects of the operation	<input type="checkbox"/> 2 Sometimes (about half the time) demonstrates familiarity with all aspects of the operation	<input type="checkbox"/> 3 Most time (more than half the time) demonstrates familiarity with all aspects of the operation	<input type="checkbox"/> 4 Almost always or always demonstrates familiarity with all aspects of the operation	Not observed

Overall, is the trainee competent to safely perform this task unsupervised: NO YES

Comments: