

## EDUCATION

# Establishing cutoff scores on assessments of surgical skills to determine surgical competence

J. Eric Jelovsek, MD; Mark D. Walters, MD; Abner Korn, MD; Christopher Klingele, MD; Nikki Zite, MD, MPH; Beri Ridgeway, MD; Matthew D. Barber, MD, MHS

**OBJECTIVE:** The aim of this study was to establish minimum cutoff scores on intraoperative assessments of surgical skills to determine surgical competence for vaginal hysterectomy.

**STUDY DESIGN:** Two surgical rating scales, the Global Rating Scale of Operative Performance and the Vaginal Surgical Skills Index, were used to evaluate trainees while performing vaginal hysterectomy. Cutoff scores were determined using the Modified Angoff method.

**RESULTS:** Two hundred twelve evaluations were analyzed on 76 surgeries performed by 27 trainees. Trainees were considered minimally competent to perform vaginal hysterectomy if total absolute scores (95% confidence interval) on Global Rating Scale = 18 (16.5–20.3)

and Vaginal Surgical Skills Index = 32 (27.7–35.5). On average, trainees met new cutoffs after performing 21 and 27 vaginal hysterectomies, respectively. With the new cutoffs applied to the same cohort of fourth-year obstetrics and gynecology trainees, all residents achieved competency in performing vaginal hysterectomy by the end of their gynecology rotations.

**CONCLUSION:** Standard-setting methods using cutoff scores may be used to establish competence in vaginal surgery.

**Key words:** assessment, competence, global assessment of operative performance, vaginal hysterectomy, vaginal surgery, vaginal surgical skills index

Cite this article as: Jelovsek JE, Walters 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.

To carry out their charge of protecting the public, licensing and certifying organizations must develop and administer assessment instruments that distinguish between trainees with ade-

quate and inadequate levels of knowledge and skill.<sup>1</sup> Unfortunately, surgical skills in obstetrics and gynecology are not directly assessed by these organizations. Although the American Board of Obstetrics and Gynecology requires a trainee to pass a written and oral examination that allows them to demonstrate they “know how” to perform surgical skills, it assumes that trainees can appropriately “perform” these skills. There is no requirement for these organizations to directly observe performance; rather, direct observation during surgery occurs by supervising surgeons during residency or fellowship. Trainees are often deemed competent to perform procedures based on nonvalidated, subjective or objective global assessments and case logs.<sup>2</sup> However, these forms of assessment do not provide direct evidence that an individual competently performs a skill, because case logs lack content validity and most forms of global assessment have poor reliability and unknown validity.<sup>2</sup>

Therefore, the challenge continues for surgeon educators to establish more robust methods of determining surgical competency that truly reflect safe and ef-

fective surgical care. Competency-based learning begins with setting standards. There are several methods of standard setting and they can be broadly categorized as relative (norm referencing) or absolute (criterion referencing). Norm referencing describes an individual's performance relative to his or her position within a group. For example, a resident is judged by comparison to the scores achieved by his or her resident colleagues on the same test. Although this is the most common method of referencing, it aims to rank trainees and allows trainees to be compared with one another. However, it cannot provide a clear assessment of the trainee's abilities, because there will always be a fixed number who fail. Moreover, norm referencing encourages competition, not cooperation, and it is somewhat unstable, as it will shift according to the performance of the norm group.<sup>3</sup> Criterion referencing provides a clear definition of what the trainee should be able to do and it provides a given standard to indicate a competent level of performance. Criterion referencing is also more responsive to the subject matter being taught and allows the teacher and learner to clearly pinpoint capabilities.

From the Obstetrics, Gynecology, and Women's Health Institute, Cleveland Clinic (Drs Jelovsek, Walters, Zite, Ridgeway, and Barber), 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 Meeting of the Society of Gynecologic Surgeons, New Orleans, LA, March 30–April 1, 2009.

Received Aug. 13, 2009; revised Oct. 5, 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).

This study was supported 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.2010.01.073

A common procedure for obstetrics and gynecology trainees to learn and become competent performing is hysterectomy. One in 5 hysterectomies in the United States is performed using the vaginal approach.<sup>4</sup> Vaginal surgery is considered the approach of choice for most patients requiring hysterectomy, because morbidity appears to be lower with the vaginal approach than with any other method.<sup>5,6</sup> Therefore, it is imperative for obstetrics and gynecology trainees to be competent performing vaginal hysterectomy. Our group recently determined the validity and reliability of 2 assessment scales that resident and fellow training programs may use to objectively assess intraoperative vaginal surgical skills: the Global Rating Scale (GRS) of operative performance, developed by Reznick et al,<sup>7</sup> and the Vaginal Surgical Skills Index (VSSI), specifically designed for the evaluation of vaginal surgical skills.<sup>8</sup> The primary aim of this study was to use formal standard-setting techniques to establish credible and defensible minimum cutoff scores on these intraoperative assessments of surgical skills to determine competence performing vaginal hysterectomy.

## MATERIALS AND METHODS

Institutional review board exemption was obtained, as this project involved the use of educational tests and did not affect the clinical course of patients. This is a supplemental study to our original study in which we determined the reliability and validity of 2 scales that can be used to assess trainees performing vaginal surgery: the new VSSI and the GRS of operative performance. To summarize our original study, the GRS was developed by Reznick et al<sup>7</sup> and consists of a 7-item global rating scale that allows supervising surgeons to directly rate important but generic skills during surgical procedures. The GRS is currently the only recommended assessment instrument listed in the Accreditation Council for Graduate Medical Education's Assessment Toolbox for assessing surgical skills. This instrument has good internal consistency (Cronbach's  $\alpha = 0.95$ ) and acceptable intrarater reliability (in-

tra-class correlation coefficient [ICC], 0.64) but low interrater reliability (ICC, 0.31) for assessing a trainee's surgical skills while performing vaginal surgery. The VSSI consists of 13 surgical principles. The VSSI also has good internal consistency (Cronbach's  $\alpha = 0.96$ ), acceptable intrarater reliability (ICC, 0.82), and better interrater reliability (ICC, 0.53) than the GRS during the assessment of trainees performing live vaginal surgery. Both instruments appear to be valid assessment scales for vaginal surgery, because they demonstrate convergent validity (how closely a new scale is related to other measures of the same construct to which it should be related) and discriminate validity (ability to distinguish between training levels).

The GRS (range, 0–35) and VSSI (range, 0–52) were used to evaluate postgraduate trainees while performing vaginal hysterectomy in obstetrics and gynecology from 2 academic medical centers. A 10-cm visual analog scale (VAS) indicating a trainee's overall level of surgical performance was also completed. A higher score on all scales indicates better performance. Vaginal hysterectomies were performed by obstetrics and gynecology residents and Female Pelvic Medicine and Reconstructive Surgery fellows (postgraduate years [PGY], 1–7) between May 2007–June 2008 and were observed live and video recorded to validate a new intraoperative surgical scale for vaginal surgery, the VSSI. Trainees performed a vaginal hysterectomy while the procedure was videotaped in a blinded, standardized fashion. An expert surgeon scored the trainee using all 3 assessment scales immediately after the procedure and again 4 weeks after the procedure using the videotape. A second blinded surgeon at a third participating institution evaluated all the videotapes using the same scales.

Because the aim of this study was to establish credible and defensible minimum cutoff scores on these 2 valid and reliable intraoperative assessments of surgical skills, methods of credible standard setting and procedures for establishing defensible absolute passing scores on performance examinations in health profession education were applied to the

assessment scales.<sup>1,9</sup> Content expert surgeons included 7 gynecologic surgeons representing the East, South, Midwest, and West Coast of the United States from 3 different academic medical centers, including: 4 from the Cleveland Clinic, 1 from Mayo Clinic, 1 who recently located to the Cleveland area from the University of Tennessee at Knoxville, and 1 surgeon from the University of California, San Francisco. Two of 7 experts were women, all were board-certified obstetricians/gynecologists, and 4 of the 7 finished an American Board of Obstetrics and Gynecology/American Urological Association-approved fellowship in Female Pelvic Medicine and Reconstructive Surgery. All experts were familiar with assessment tools, curricula, and trainees. Historically, most standard-setting studies have demonstrated that judges, absent all performance data, tend to set unrealistically high passing scores, which will fail an unreasonably high proportion of trainees. To avoid this, it is recommended that experts are "calibrated" for the standard-setting methods to have a realistic expectation of actual trainee performance by first-hand observation of the actual scale scores for trainees and in-depth discussion of differences between competence and expertise.<sup>9,10</sup> This was performed for this study.

Most absolute standard-setting methods are based around the concept of a borderline trainee's performance. The concept of a borderline trainee originated from Angoff and assumes that a borderline trainee is one who has an exactly 50:50 probability of passing or failing the assessment.<sup>10</sup> Therefore, a minimum cutoff score is 1 that separates those who are competent and those who are not. For this study, cutoff scores were determined using a single method, the Modified Angoff method, and confirmed using scores derived from 2 additional standard-setting methods, the Contrasting Groups method<sup>11,12</sup> and Hofstee method.<sup>10,13,14</sup> Each of these standard-setting methods is described later.

In this study, the Modified Angoff method began as the experts discussed the characteristics and gave examples of

Download English Version:

<https://daneshyari.com/en/article/3435336>

Download Persian Version:

<https://daneshyari.com/article/3435336>

[Daneshyari.com](https://daneshyari.com)