Development and validation of a laparoscopic hysterectomy cuff closure simulation model for surgical training

Bitton, E; Propst, K; Muffly, T.
AJOG March 2016.
Background

• TLH volume is decreasing
• Vaginal cuff dehiscence occurs higher in laparoscopic approaches
• Simulation based training can fill a gap in surgical skill development
• Prior to operating on people – trainees should be competent in abilities
• Can we make a model that is a valid and reliable as an assessment and training tool?
Methods

• FLS box trainer
• RUMI + the crafting closet
• Participants: 2 academic centers trainees and teachers + attendees at IUGA/AUGS
• 2 groups: “Expert” and “Trainee”

FIGURE
Laparoscopic vaginal cuff closure simulation model

A, RUMI Advanced Uterine Manipulation System (Cooper Surgical) attached to the FLS box trainer
B, Simulated vaginal cuff made of the neoprene and swimsuit material and vaginal manipulator stem
C, Custom bracket attachment. D, View of the simulator in the FLS box.
Methods

Expert

• Self-identify
• “perform laparoscopic gynecologic surgery routinely as part of their practice and instruct residents and/or fellows”

Trainees

• Senior residents and fellows in PMRS
Methods

Recruit at AUGS/IUGA

Self Identify as expert/trainee

Procedure
- 2 stitches
- 3 extracorporeal throws

Blinded review of silent video using GOALS by 2 experts
GOALS scale

- Global Operative Assessment of Laparoscopic Skills
- 5 domains: depth perception, bimanual dexterity, efficiency, tissue handling, and autonomy.
- 5 points per domain
- Modified for this study:
  - needle handling
  - knot tying
  - incorporation of the epithelial edge
Methods: Analysis

• Student T Test to compare scores and operating times
• 90% power to detect difference of 7 points between the two groups = alpha of 0.05
Results

- Great variety in experience
- No trainees prefer extracorporeal knot tying
- “Construct validity”-experts tied faster and scored higher on the “GOALS”
- 85% of participants agree or strongly agree that model is realistic
- All agree model is helpful for improving technique
- $100-200 is acceptable cost
- The blinded experts agreed 80% of the time

<table>
<thead>
<tr>
<th>TABLE 2</th>
<th>Participant demographics and surgical experience</th>
</tr>
</thead>
<tbody>
<tr>
<td>Variable</td>
<td>Experts (n = 19)</td>
</tr>
<tr>
<td>Age, y</td>
<td>39 (34-50)</td>
</tr>
<tr>
<td>Male/female</td>
<td>7/12</td>
</tr>
<tr>
<td>Right/left handed</td>
<td>17/1</td>
</tr>
<tr>
<td>Completed FPMRS fellowship</td>
<td>13 (68.4%)</td>
</tr>
<tr>
<td>Completed gynecology/oncology fellowship</td>
<td>3 (15.8%)</td>
</tr>
<tr>
<td>Surgical experience beyond training, y</td>
<td>12 (3-23)</td>
</tr>
<tr>
<td>Laparoscopic vaginal cuff closure (total since residency)</td>
<td>55 (10 to &gt; 150)</td>
</tr>
<tr>
<td>Procedures requiring laparoscopic suturing (total since residency)</td>
<td>60 (10 to &gt; 150)</td>
</tr>
<tr>
<td>Laparoscopic vaginal cuff closures in 1 y</td>
<td>4 (0 to &gt; 20)</td>
</tr>
<tr>
<td>Procedures requiring laparoscopic suturing in 1 y</td>
<td>8 (0 to &gt; 20)</td>
</tr>
<tr>
<td>Laparoscopic/robotic vaginal cuff closures in 1 y</td>
<td>15 (0 to &gt; 20)</td>
</tr>
<tr>
<td>Frequency of extracorporeal/intracorporeal knot tying</td>
<td>13 (68.4%)/3 (15.8%)</td>
</tr>
<tr>
<td>Preferred method of vaginal cuff closure</td>
<td></td>
</tr>
<tr>
<td>Intracorporeal</td>
<td>0 (0%)</td>
</tr>
<tr>
<td>Extracorporeal</td>
<td>10 (52.3%)</td>
</tr>
<tr>
<td>Vaginal closure</td>
<td>3 (15.8%)</td>
</tr>
<tr>
<td>Robotically assisted</td>
<td>6 (31.6%)</td>
</tr>
</tbody>
</table>
### Results

**TABLE 3**

**Construct validity: time (seconds) and performance (median score [range]) on the simulation model**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Experts (n = 19)</th>
<th>Trainees (n = 21)</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Time, s</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Needle loading</td>
<td>14 (3–60)</td>
<td>18 (8–80)</td>
<td>.95</td>
</tr>
<tr>
<td>Stitch placement</td>
<td>27 (13–82)</td>
<td>43 (11–161)</td>
<td>.01</td>
</tr>
<tr>
<td>Knot tying (x4)</td>
<td>54 (29–229)</td>
<td>112 (42–250)</td>
<td>.003</td>
</tr>
<tr>
<td><strong>GOALS metrics</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Depth perception</td>
<td>4.5 (2–5)</td>
<td>3 (2–4.5)</td>
<td>.001</td>
</tr>
<tr>
<td>Bimanual dexterity</td>
<td>5 (2–5)</td>
<td>2.5 (1.5–4)</td>
<td>.001</td>
</tr>
<tr>
<td>Efficiency</td>
<td>4.5 (3–5)</td>
<td>2.5 (2–4.5)</td>
<td>.001</td>
</tr>
<tr>
<td>Tissue handling</td>
<td>5 (2–5)</td>
<td>2 (1.5–5)</td>
<td>.001</td>
</tr>
<tr>
<td>GOALS total score (20 points)</td>
<td>18.8 (11–20)</td>
<td>10 (8–18)</td>
<td>.001</td>
</tr>
<tr>
<td><strong>Novel metrics</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Needle handling</td>
<td>5 (2–5)</td>
<td>2 (2–4)</td>
<td>.001</td>
</tr>
<tr>
<td>Knot tying</td>
<td>4.75 (2–5)</td>
<td>2.5 (1.5–4.5)</td>
<td>.001</td>
</tr>
<tr>
<td>Incorporation of mucosa</td>
<td>5 (3–5)</td>
<td>3 (2–5)</td>
<td>.001</td>
</tr>
<tr>
<td>Total score (35 points)</td>
<td>33 (18–35)</td>
<td>17.5 (14–31.5)</td>
<td>.001</td>
</tr>
</tbody>
</table>

Each metric is scored from 1 to 5, with a maximum total GOALS score of 20 (autonomy category was omitted). Combined with 3 novel metrics, the maximum possible score is 35 points. The median score was calculated between the mean score assigned by 2 blinded experts. Interclass correlation coefficient for the GOALS score and the total score was 0.80 (95% confidence interval, 0.70–0.93). GOALS, Global Overall Operative Assessment of Laparoscopic Skills.

Conclusions/Comment

• Simulation is important
• Laparoscopic hysterectomies have higher dehiscence of vaginal cuff
• This model was easy to make
• “For this model we have demonstrated evidence of construct validity using a valid and reliable assessment scale”
Evaluating The Literature:
Validation Research in GYN Surgery
Simulation & Assessment

Edgar L LeClaire M.D.
Take home points...

• **How** we must validate is as important as **why** we must validate.

• Validation research [should] focus on 5 factors that impact **score meaning**: includes rating scales filled out by an observer or a simulator with an embedded scoring system.

• Compliance with validation standards is needed to develop **trustworthy and meaningful** competency scores.
Surgical Learning Environment

**Encounters**
- Personnel
- Diagnostics
- Intervention
- Goal (confirmed by diagnostics)

**Clinical**
- Patient
- Lab Test
- Treatment
- Health

**Training**
- Resident
- Assessment
- Training
- Competency
“See one, do one, teach one.”
GYN Education Challenges

- Increasing surgical modalities (and tech therein)
  - Laparoscopic, Robot-assisted, Open, Endoscopic, and Vaginal
- Duty hour restrictions, patient volume changes
  - Less time to teach, fewer learning opportunities
- Increased subspecialization
  - Decreased resident:faculty one-on-one time
- Growing emphasis on patient safety and outcomes-based reimbursement
  - Payers, hospitals, specialty boards, and government
The Importance of Validity

• Improved safety and quality outcomes
  – requires...
• Safe and independent practice
  – requires...
• Verification of skills acquisition
  – requires...
• Assessment Validity
• Validity Standards
# GOALS
Global Operative Assessment of Laparoscopic Skills

<table>
<thead>
<tr>
<th>Domains</th>
<th>Anchor Descriptors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Depth Perception</td>
<td>Constantly overshooting target, hits backstop, wide swings, slow to correct</td>
</tr>
<tr>
<td></td>
<td>Some overshooting or missing plane but corrects quickly</td>
</tr>
<tr>
<td></td>
<td>Accurately directs instruments in correct plane to target</td>
</tr>
<tr>
<td>Bimanual Dexterity</td>
<td>Use of one hand, ignoring nondominant hand, poor coordination between hands</td>
</tr>
<tr>
<td></td>
<td>Use of both hands but does not optimize interactions between hands to facilitate conduct of operation</td>
</tr>
<tr>
<td></td>
<td>Expertly uses both hands in a complementary manner to provide optimal working exposure</td>
</tr>
<tr>
<td>Efficiency</td>
<td>Uncertain, much wasted effort, many tentative motions, constantly changing focus of operation, or persisting at a task without progress</td>
</tr>
<tr>
<td></td>
<td>Slow, but planned and reasonably organized</td>
</tr>
<tr>
<td></td>
<td>Confident, efficient and safe conduct of operation, maintaining focus on component of procedure until better done by another approach</td>
</tr>
<tr>
<td>Tissue Handling</td>
<td>Rough, tears tissue by excessive traction, injures adjacent structures, poor control of coagulation device (recoil), grasper frequently slips off</td>
</tr>
<tr>
<td></td>
<td>Handles tissues reasonably well, with some minor trauma to adjacent tissues, eg, coagulation of liver, causes unnecessary liver bleeding, occasional slipping of grasper</td>
</tr>
<tr>
<td></td>
<td>Handles tissues very well with appropriate traction on tissues and negligible injury of adjacent structures. Uses energy sources appropriately but not excessively</td>
</tr>
<tr>
<td>Autonomy</td>
<td>Unable to complete entire procedure, even in a straightforward case and with extensive verbal guidance</td>
</tr>
<tr>
<td></td>
<td>Able to complete operation safely with moderate prompting</td>
</tr>
<tr>
<td></td>
<td>Able to complete operation independently without prompting</td>
</tr>
</tbody>
</table>

Adapted from Vassiliou et al, American Journal of Surgery, with permission.
Thought experiment...

- The American Congress of Obstetrics and Gynecology implements a standardized simulation-based assessment candidates **must pass** in order to obtain certification.
Validation Standards

- Developers
- Consequences / Fairness
- Reliability
- Relations to other Variables
- Scoring System

1999 Standards for Educational & Psychological Testing, APA, AERA, NCME
Question 1:  
What sort of things about the authors / developers of this high-stakes assessment would you want to know?
Question 1:
What sort of things about the authors / developers of this high-stakes assessment would you want to know?

• How many?
• Qualifications?
• Years experience in the surgical domain?
• Are they recognized as experts by their peers?

• P.S. How should we define surgical expertise?
Validation Standards

Developers

Consequences / Fairness

Reliability

Validity

Relations to other Variables

Scoring System

1999 Standards for Educational & Psychological Testing, APA, AERA, NCME
Question 2:
Is it reliable?
Question 2: Is it reliable?

- If a scoring rubric or rating scale
  - Interrater reliability
- If an embedded scoring system within a simulator
  - Reliability across user types
Validation Standards

Developers

Consequences / Fairness

Reliability

Validity

Relations to other Variables

Scoring System

1999 Standards for Educational & Psychological Testing, APA, AERA, NCME
Question 3:
What sort of things about the grading system would you want to know?
Question 3:
What sort of things about the grading system would you want to know?

• Does it make clinical sense?
  – Will you fail if you fail to recognize hemorrhage? If so at what volume?
  – Is there a time component?

• In short, does the scoring model reward best practices and punish dangerous errors appropriately?

• Probability of “accidentally” passing test?
Validation Standards

Developers

Validity

Consequences / Fairness

Reliability

Relations to other Variables

Scoring System

1999 Standards for Educational & Psychological Testing, APA, AERA, NCME
Validation Standards

Developers
Consequences / Fairness
Validity
Relations to other Variables
Scoring System
Reliability

1999 Standards for Educational & Psychological Testing, APA, AERA, NCME
Validation Standards

- Test-Criterion: clinical outcome
- Generalizability: novice vs. expert, within different settings and venues
- Convergence: scores from other assessments
Question 4:
How do scores from this assessment relate to other factors?

- to scores from other OB/GYN skills assessments?
- to the skill level of the test-taker (novice / low-intermediate / high-intermediate / expert)?
- to actual, relevant surgical outcomes from the real world?
Question 4:
How do scores from this assessment relate to other factors?

• to scores from other OB/GYN skills assessments?
  – should correlate positively

• to the skill level of the test-taker (novice / intermediate / expert)?
  – should predict score achievement

• to actual, relevant surgical outcomes?
  – should correlate positively with desired outcomes (e.g. decreased EBL)
  – should correlate negatively with undesired outcomes (e.g. bleeding complications, etc.)
Validation Standards

- Test-Criterion: clinical outcome
- Generalizability: novice vs. expert, within different settings and venues
- Convergence: scores from other assessments

1999 Standards for Educational & Psychological Testing, APA, AERA, NCME
Validation Standards

1999 Standards for Educational & Psychological Testing, APA, AERA, NCME
Question 5:
What have been the consequences of testing (good and bad)?
Question 5:
What have been the consequences of testing (good and bad)?

• Positively impacted patient safety outcomes?
• Negatively impacted providers in an unforeseen way?
• Are score-based decisions (“to certify or not to certify”) justified by the evidence?
• Is it fair?
Validation Standards

- Developers
- Consequences / Fairness
- Reliability
- Relations to other Variables
- Scoring System

1999 Standards for Educational & Psychological Testing, APA, AERA, NCME
Samuel Messick

“...scientific inquiry into score meaning.”
STANDARDS
for educational and psychological testing

American Educational Research Association
American Psychological Association
National Council on Measurement in Education
Validity: applying current concepts and standards to gynecologic surgery performance assessments

Edgar L. LeClaire · Mikio A. Nihira · Patricia L. Hardré

Received: 19 November 2013/Accepted: 19 August 2014/Published online: 22 November 2014
© Springer Science+Business Media Dordrecht 2014

Abstract Validity is critical for meaningful assessment of surgical competency. According to the Standards for Educational and Psychological Testing, validation involves the integration of data from well-defined classifications of evidence. In the authoritative framework, data from all classifications support construct validity claims. The two aims of this study were to develop a categorization method for validity evidence published in support of surgery performance assessments and to summarize the results of applying this methodology to the gynecologic surgery literature. This was a critical analysis of published observations reported as validity evidence in studies with a construct validity claim. Medline and Embase databases were searched using keywords: “surgery” and “construct validity”. Parameters included English-language articles published from 2000 to 2012. Gynecologic studies were analyzed for definitions of construct validity and nonstandard terminology. Categorization criteria were developed and applied by the researchers to all observations. Two independent evaluators examined reported observations for compliance with guidelines provided by the Standards. Inter-rater agreement was calculated using weighted kappa. The initial search returned 167 articles. Twenty-five articles were left for inclusion in our analysis. Eighteen (72 %) articles defined construct validity as the ability to discriminate between expert and novice levels of proficiency. Within the sample, 80

Electronic supplementary material The online version of this article (doi:10.1007/s10459-014-9548-y) contains supplementary material, which is available to authorized users.

E. L. LeClaire (✉)
Division of Female Pelvic Medicine and Reconstructive Surgery, Department of Obstetrics and Gynecology, University of Kansas School of Medicine—Wichita, 551 North Hillside, Suite 500,
“Validated” Assessment Tools in GYN Surgery

Table 2. Performance measurement instruments (simulators, scoring rubrics, and other tests) identified in the sample \( (n=25) \)

<table>
<thead>
<tr>
<th>Instrument</th>
<th>Number of validation studies</th>
</tr>
</thead>
<tbody>
<tr>
<td>Essure Sim VR (Simbionix, Cleveland, OH)</td>
<td>1</td>
</tr>
<tr>
<td>Interactive Voice Response Instrument</td>
<td>1</td>
</tr>
<tr>
<td>Lap Sim-Gyn (Surgical Science Ltd, Gothenburg, Sweden)</td>
<td>2</td>
</tr>
<tr>
<td>Laparoscopic box trainer tasks*</td>
<td>5</td>
</tr>
<tr>
<td>Laparoscopic Skills Testing and Training Model (European Academy of Gynaecological Surgery)</td>
<td>1</td>
</tr>
<tr>
<td>Objective Assessment of Surgical Competence in Gynaecological Laparoscopy</td>
<td>1</td>
</tr>
<tr>
<td>Objective Structured Assessment of Technical Skills*</td>
<td>9</td>
</tr>
<tr>
<td>Objective Structured Clinical Examination for Colposcopy</td>
<td>1</td>
</tr>
<tr>
<td>Postgraduate Obstetrics and Gynaecology Progress Test</td>
<td>1</td>
</tr>
<tr>
<td>Script Concordance Test</td>
<td>1</td>
</tr>
<tr>
<td>Surgical Sim (METI, Sarasota, FL)</td>
<td>1</td>
</tr>
<tr>
<td>Vaginal Surgery Skills Index</td>
<td>1</td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td><strong>25</strong></td>
</tr>
</tbody>
</table>

\*Collectively grouped instrument type
Gaps in Validity Evidence

Figure 2. Frequency of Standard classifications among all discrete observations (n=155) reported as validity evidence within sample

Legend: C = Content, RP = Response Process, IS = Internal Structure, GEN = Generalizability, TCR = Test-Criterion, CONV = Convergent, DISC = Discriminant, CONSQ = Consequential
Development and validation of a laparoscopic hysterectomy cuff closure simulation model for surgical training

Elena Tunitsky-Bitton, MD; Katie Propst, MD; Tyler Muffly, MD

BACKGROUND: The number of robotically assisted hysterectomies is increasing, and therefore, the opportunities for trainees to become competent in performing traditional laparoscopic hysterectomy are decreasing. Simulation-based training is ideal for filling this gap in training.

OBJECTIVE: The objective of the study was to design a surgical model for training in laparoscopic vaginal cuff closure and to present evidence of its validity and reliability as an assessment and training tool.

STUDY DESIGN: Participants included gynecology staff and trainees at 2 tertiary care centers. Experienced surgeons were also recruited at the combined International Urogynecologic Association and American Urogynecologic Society scientific meeting. Participants included 19 experts and 21 trainees. All participants were recorded using the laparoscopic hysterectomy cuff closure simulation model. The model was constructed using the an advanced uterine manipulation system with a sacrocolpopexy tip/vaginal stent, a vaginal cuff constructed from neoprene material and lined with a swimsuit material (nylon and spandex) secured to the vaginal stent with a plastic cable tie. The uterine manipulation system was attached to the fundamentals of laparoscopic surgery laparoscopic training box trainer using a metal bracket. Performance was evaluated using the Global Operative Assessment of Laparoscopic Skills scale. In addition, needle handling, knot tying, and incorporation of epithelial edge

RESULTS: Total and annual experience with laparoscopic suturing and specifically vaginal cuff closure varied greatly among the participants. For the construct validity, the participants in the expert group received significantly higher scores in each of the domains of the Global Operative Assessment of Laparoscopic Skills Scale and for each of the 3 added items than did the trainees. The median total Global Operative Assessment of Laparoscopic Skills Scale score (maximum 20) for the experts was 18.8 (range, 11–20), whereas the median total Global Operative Assessment of Laparoscopic Skills Scale score for the trainees was 10 (range, 8–18) (P = .001). The overall score that included the 3 new domains (maximum 35) was 33 (range, 18–35) for the experts and 17.5 (range, 14–31.5) for trainees (P = .001). For the face validity testing, the majority of the study participants (32 [85%]) agreed or strongly agreed that the model is realistic and all participants agreed or strongly agreed that the model appears to be useful for improving technique required for this task. For the inter-rater reliability, the scores assigned by each observer had an inter-class correlation coefficient of 0.8 (95% confidence interval, 0.7–0.93).

CONCLUSION: This model is easily constructed and has an acceptable cost. We have demonstrated evidence of construct validity. This is a valuable education tool that can serve to improve skills, which are essential to the gynecological surgeon but are often lacking in residency training.
The 5 Questions of Validity

- 1: credentials of simulator designers not disclosed, experts were self-selected
- 2: inter-rater reliability of GOALS scores
- 3: grading system: not described, time only
- 4: scores were not correlated to anything
- 5: no consequences proposed or reported

Final Analysis: GOALS validity evidence, box trainer that is “useful” according to experts
## GOALS
Global Operative Assessment of Laparoscopic Skills

<table>
<thead>
<tr>
<th>Domains</th>
<th>Anchor Descriptors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Depth Perception</td>
<td>Constantly overshooting target, hits backstop, wide swings, slow to correct</td>
</tr>
<tr>
<td></td>
<td>Some overshooting or missing plane but corrects quickly</td>
</tr>
<tr>
<td></td>
<td>Accurately directs instruments in correct plane to target</td>
</tr>
<tr>
<td>Bimanual Dexterity</td>
<td>Use of one hand, ignoring nondominant hand, poor coordination between hands</td>
</tr>
<tr>
<td></td>
<td>Use of both hands but does not optimize interactions between hands to facilitate</td>
</tr>
<tr>
<td></td>
<td>conduct of operation</td>
</tr>
<tr>
<td></td>
<td>Expertly uses both hands in a complementary manner to provide optimal working exposure</td>
</tr>
<tr>
<td>Efficiency</td>
<td>Uncertain, much wasted effort, many tentative motions, constantly changing focus</td>
</tr>
<tr>
<td></td>
<td>of operation, or persisting at a task without progress</td>
</tr>
<tr>
<td></td>
<td>Slow, but planned and reasonably organized</td>
</tr>
<tr>
<td></td>
<td>Confident, efficient and safe conduct of operation, maintaining focus on component</td>
</tr>
<tr>
<td></td>
<td>of procedure until better done by another approach</td>
</tr>
<tr>
<td>Tissue Handling</td>
<td>Rough, tears tissue by excessive traction, injures adjacent structures, poor control</td>
</tr>
<tr>
<td></td>
<td>of coagulation device (recoil), grasper frequently slips off</td>
</tr>
<tr>
<td></td>
<td>Handles tissues reasonably well, with some minor trauma to adjacent tissues, eg,</td>
</tr>
<tr>
<td></td>
<td>coagulation of liver, causes unnecessary liver bleeding, occasional slipping of</td>
</tr>
<tr>
<td></td>
<td>grasper</td>
</tr>
<tr>
<td></td>
<td>Handles tissues very well with appropriate traction on tissues and negligible injury</td>
</tr>
<tr>
<td></td>
<td>of adjacent structures. Uses energy sources appropriately but not excessively</td>
</tr>
<tr>
<td>Autonomy</td>
<td>Unable to complete entire procedure, even in a straightforward case and with</td>
</tr>
<tr>
<td></td>
<td>extensive verbal guidance</td>
</tr>
<tr>
<td></td>
<td>Able to complete operation safely with moderate prompting</td>
</tr>
<tr>
<td></td>
<td>Able to complete operation independently without prompting</td>
</tr>
</tbody>
</table>

Adapted from Vassiliou et al, American Journal of Surgery, with permission.
Defining Expertise in Gynecologic Surgery: Perspectives of Expert Gynecologic Surgeons

Patricia L. Hardré, PhD,* Mikio Nihira, MD, MPH,† Edgar LeClaire, MD,‡ and Michael Moen, MD§

Objective: The aim of this study was to describe how professional expertise is defined and understood among gynecologic surgeons and what experiential factors contribute to that understanding.

Methods: Semistructured interviews with 16 experts in Female Pelvic Medicine and Reconstructive Surgery were conducted to identify how expertise in their field is defined, recognized, and assessed. Independent thematic analysis of the interview transcripts was performed by each member of the research team and then distilled and synthesized into convergent themes.

Results: Experts described surgical expertise as difficult to define but with several dominant themes including knowledge, technical skills, clinical experience, adaptability, continuous learning, communication, and professional recognition. Expertise requires judgment in applying technical skills to meet each patient's specific needs. Experts described unique ways of seeing and thinking during surgery, characterized by spatial awareness of relevant anatomy, temporal awareness of future changes, and rapidly adaptive application of their skills enabling them to do difficult tasks with fluidity, making the tasks seem easy to observers. These expert surgeons acknowledged that achieving expertise requires hard work and maintaining expertise requires continuous learning, highlighted by challenge seeking to do the most difficult tasks in their field. They also noted the importance of effective communication of their knowledge to others, which contributes to their perception as experts by colleagues in the field.

Conclusions: Surgical expertise is a complex phenomenon with several meaningful themes. Understanding the authentic nature of surgical expertise can be used to support the development of competencies and the effective mentoring of promising surgical trainees to achieve surgical expertise.

Key Words: surgical expertise, surgical skills, surgical competence

Expertise in highly technical professions is complex. Depending on the precise nature of the work, expertise can be defined in a number of ways based on the level of education, past experiences, criteria for “successful” or “expert” performance, which complicates efforts to codify it for educational and evaluation needs.

Every skill and profession has a knowledge base composed of both basic and advanced knowledge and a skill set composed of both general (cross-disciplinary) and field or task-specific (domain) skills. General skills in surgery include things such as observation, critical reasoning, analysis, problem solving, and adaptive thinking. Domain-specific skills in surgery include manual precision, accurate tool-handling, specific surgical techniques, and specific anatomical knowledge. Within a given field or domain of practice, competency is defined by a certain body of knowledge and skills, but expertise moves beyond core competencies to a level that involves doing the job faster, better, more effectively, and more adaptively and with fewer errors. These are the characteristics that distinguish experts from novices in professional practice.

Knowing that some physicians become good, solid practitioners, but with minimal recognition, whereas others develop higher levels of expertise in their field, medical education can benefit from a better understanding of expertise development. Not only does the clear definition of expertise drive educational design, it also supports development through learner knowledge of, and self-regulation toward, clear performance targets. Because of the specialized nature of professional expertise, the best initial source to understand it is the experts themselves. The objective of this study was to identify specifically how expert gynecologic surgeons articulate the definitions and standards of excellence in their professional field. This understanding can be used to improve current physician training and competence assessment, contributing to the development of expertise.

MATERIALS AND METHODS

This study was approved by the institutional review board...