



## Evaluation of obstetrics procedure competency of family medicine residents

Haijun Wang<sup>1</sup>, Eric Warwick<sup>1</sup>, Maria C. Mejia de Grubb<sup>1</sup>, Nanfu Deng<sup>2</sup>, Jane Corboy<sup>1,3</sup>

### Abstract

**Objective:** To establish a procedure evaluation system to monitor residents' improvement in obstetrics (OB) procedures performance and skills during the training period.

**Methods:** A web-based procedure logging and evaluation system was developed using Microsoft.net technology with a SQL server as a backend database. Residents' logged OB procedures were captured by the system. The OB procedures logged within 7 days were evaluated by supervising faculty using three observable outcomes (procedure competency, procedure-related medical knowledge level, and patient care).

**Results:** Between 1 July 2005 and 30 June 2008, a total of 8543 procedures were reported, of which 1263 OB procedures were evaluated by supervising faculty. There were significant variations in the number of logged procedures by gender, residency track, and US versus non-US medical graduates. Approximately 84% of the procedures were performed (independently or with assistance) by residents. Residents' procedure skills, procedure-related medical knowledge, and patient care skills improved over time, with significant variations by gender among the three outcomes.

**Conclusion:** The benefits of specific evaluation of procedural competence in postgraduate medical education are well established. Innovative and reliable tools to assess and monitor residents' procedural skills are warranted.

**Keywords:** Resident procedure; family medicine; evaluation

1. Department of Family and Community Medicine, Baylor College of Medicine, Houston, TX, USA

2. Medical School, Baylor College of Medicine, Houston, TX, USA

3. Department of Family Medicine, Oregon Health & Science University, 3181 SW Sam Jackson Park Rd, Portland, OR 97239-3098, USA

### CORRESPONDING AUTHOR:

Haijun Wang, PhD, MPH  
Department of Family and Community Medicine, 3701 Kirby Drive, Suite 600, Houston, TX 77098, USA

Tel.: +713-798-7744

E-mail: haijunw@bcm.edu

Received 9 April 2015;

Accepted 11 May 2015

### Introduction

Most agree that tracking the number of procedures residents perform alone does not satisfy proof of competency [1–3]. Evaluation of residents' procedural competencies is a critical component of family medicine residency training by appropriately assessing resident's abilities in procedural training in preparation for a successful practice [2, 4].

The Family Medicine Milestones project, a joint initiative of the Accreditation Council for Graduate Medical Education (ACGME) and the

American Board of Family Medicine, postulates the need for a framework to assess the development of resident competencies in key areas, including evaluation of procedures and clinical experience in preparation for the newly implemented Graduate Medical Education (GME) accreditation system [5]. This system aims to prepare physicians to practice in the 21st century on the basis of educational outcomes [6]; however, this system requires innovative approaches and expected measurable or observable attributes



## Procedure evaluation of family medicine residents

(knowledge, abilities, skills, or attitudes) of residents at key stages in their training.

When affirming post-residency procedural privileges, program directors rely primarily on procedure count and observation. Although specialty-appropriate procedures, including obstetrics (OB) procedures, have been the scope of training in family medicine residencies, a consensus statement, including a core list of standard procedures that all family medicine residents should be able to perform by the time of graduation, has been previously published [7, 8]. Nevertheless, documentation of procedural competency is a challenging issue.

Furthermore, there are technical and cognitive aspects to each procedure that must be mastered for the clinician to be deemed competent. It is difficult to predict how many times a resident must repeat a procedure to achieve competency. Attempts to estimate those numbers have been made [9], but these estimates have been shown to be inaccurate for some procedures, as some residents required more than twice the recommended number to achieve competency [10]. Although other methods have been proposed, including computer simulation, knowledge testing, and the use of cadavers, the most common method of determining competency remains direct supervision and certification of the resident by an experienced and privileged attending physician [2].

The objective of this study was to establish a procedure evaluation system to monitor residents' improvement specially for OB procedures during the training period as part of continuous efforts to enhance qualitative measures to assess the development of the resident competencies.

### Methods

The Institutional Review Board of Baylor College of Medicine approved this study.

The procedure log and evaluation system was developed using Microsoft ASP.NET with a Microsoft SQL server 2000 as a backend database. Although quantifiable procedure data were collected since 2002, qualitative data, such as how well residents performed procedures, was not recorded until later. To address the missing data, a qualitative procedure evaluation tool based on three observable outcomes was added to

this system in 2005. The flowchart of the system is shown in Fig. 1.

The pertinent data for each OB procedure reported by residents from 1 July 2005 to 30 June 2008 included basic information (date of procedure, patient name, date of birth, gender, and medical record number), as well as procedure-related information (procedure category, procedure name, level of involvement, and supervisor). To ensure the timeliness of procedure evaluation, only cases reported within 7 days post-procedure were evaluated. If the procedure was related to labor and delivery, continuity delivery questions were added. Continuity OB delivery is defined as a series of encounters with a pregnant patient that includes recurrent prenatal evaluations and management of labor and delivery. Continuity OB delivery also includes infant and mother follow-up through postnatal/postpartum care [11]. On the day following the procedure, an automated email messenger system reminded the supervising faculty to evaluate the resident's performance from the previous day. OB procedures were evaluated with the following three criteria: procedural competency; medical knowledge; and patient management. When a continuity OB delivery was reported, the following four additional criteria were applied: general bedside manner during labor (Patient care – Professionalism); attention to psychosocial needs (Interpersonal Communications -Professionalism); communication with the patient and her family regarding the progress of labor (Interpersonal Communication); and providing postpartum care for the mother and newborn to the fullest extent possible (Patient Care). Details of these criteria are outlined in Tables 1 and 2.

Data on faculty assistance or resident independence during OB procedures were also collected. Faculty certified the procedure by selecting "yes" in the independency field. After the procedure was certified, an automated email message was sent to all faculty members privileged in the specified area (i.e., family medicine [FM] faculty supervising OB procedures). If another faculty member disagreed with the certification of competency for independent performance, the faculty member was permitted to express their concerns through the credentialing form within 4 days. Otherwise the resident was officially certified for the procedure (Fig. 1).

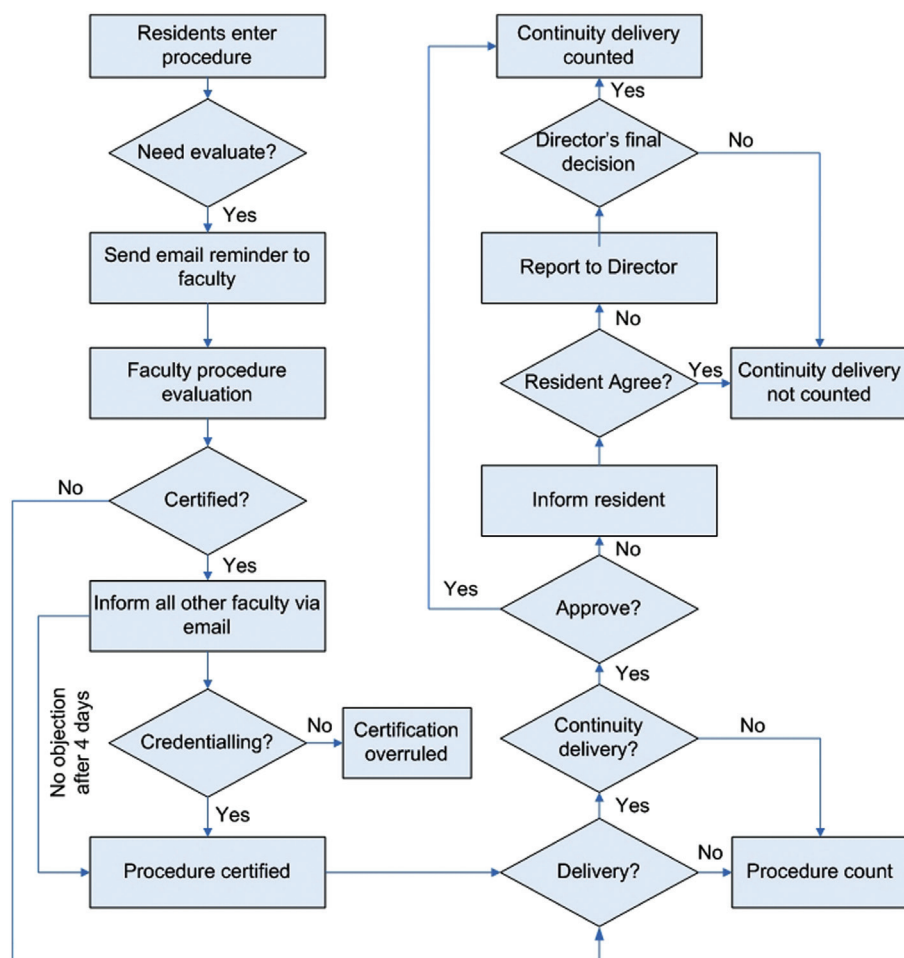


Fig. 1. Flowchart of procedure logging and evaluation system.

Continuity delivery certification is another feature of our procedure evaluation system. If a continuity procedure report made by a resident was denied by a faculty member, the resident was immediately notified of the denial via email, at which time the resident could dispute the certification denial with the program director.

### Data analysis

Descriptive statistics (frequencies, percentages, means, and standard deviations) were calculated based on the resident postgraduate year (PGY) and demographic variables. Cronbach's alpha was used to reflect the reliability of the

procedure evaluation form. Differences between cohorts were detected by t-test, one-way analysis of variance (ANOVA), and two-way ANOVA, and trend analysis was used to compare those differences. IBM SPSS software (version 22.0; Chicago, IL, USA) was used for all data analyses.

### Results

Sixty-one residents participated in the study (Table 3). The average age of residents was 30.9 years (median age, 29 years) and 62.3% were female. Self-reported ethnicities were 41.0% Asian, 26.2% African American, 16.4% Caucasian, and 16.4% Hispanic. Most residents were from non-US medical schools



Table 1. Procedure evaluation criteria

Procedural competency	<ol style="list-style-type: none"> <li>1. Cannot outline the steps of the procedure</li> <li>2. Can describe what to do, but cannot do some of the steps or does them out of cases</li> <li>3. Able to do each step with prompting of what to do next;</li> <li>4. Performs all steps in correct sequence, but hesitant or tentative</li> <li>5. Smooth performance of all steps without prompting, in “usual” situations</li> <li>6. Able to adapt to common variations in patient situations</li> </ol>
Medical Knowledge	<ol style="list-style-type: none"> <li>1. Has a gap in basic knowledge and understanding</li> <li>2. Can apply knowledge to routine and novel situations</li> <li>3. Can explain and defend decisions and judgments</li> </ol>
Patient management	<ol style="list-style-type: none"> <li>1. Deficiencies in judgment and/or technical performance</li> <li>2. Competent judgment and/or technical performance with few minor errors</li> <li>3. Proficient performance without hesitation or errors</li> </ol>

Table 2. Continuity delivery evaluation criteria

Presence	<ol style="list-style-type: none"> <li>1. Arrives just in time to deliver baby; does not participate in decisions about management of labor; does not coach patient in second stage</li> <li>2. Involved in second stage and delivery; in phone contact with on-call resident, discusses progress; involved in labor decision-making</li> <li>3. Actively involved in the management of patient’s active labor; performs necessary interventions; at bedside for all critical events</li> </ol>
Attention	<ol style="list-style-type: none"> <li>1. Is distant from patient and family, does not establish rapport or address patient needs</li> <li>2. Is attentive to patient and family; advocates for patient needs with nursing staff if necessary</li> <li>3. Elicits and respects values; culturally sensitive; empathetic; ethical; sensitive to confidentiality</li> </ol>
Communication	<ol style="list-style-type: none"> <li>1. Gives incomplete or inaccurate information to patient; coaches ineffectively; gives unclear picture of situation to family, patient, or attending</li> <li>2. Adequately conveys accurate and complete information; adequate labor coach; explains why procedures are necessary</li> <li>3. Comfortably conveys information fully and accurately; verifies understanding; delivers “bad news” about labor progress empathically and clearly</li> </ol>
Newborn care	<ol style="list-style-type: none"> <li>1. Does not complete newborn exam; does not write postpartum note; does not check out patient to on-call resident</li> <li>2. Completes newborn exam; writes postpartum note; checks out patient to on-call resident</li> <li>3. Sees mom and baby at least once postpartum; facilitates and encourages bonding and breast-feeding</li> </ol>

and only 37.7% of residents were US medical school graduates. Approximately 30% of residents began their within 1 year of medical school graduation, while 14.8% started residency more than 10 years post-graduation (all non-US medical school graduates). During this training period, the residency program consisted of two tracks (traditional practice and

urban underserved). The intent of the traditional practice track is to develop leaders in the delivery of comprehensive, quality health care services through multispecialty practice. The mission of the urban underserved track centers on using the biopsychosocial model to improve health for underserved communities. Except for the hiatus between residency and



Table 3. Resident demographic information

Demographics	US Graduate	Non-US graduate	Total
PGY*			
PGY-1	14(12.3%)	21(18.3%)	36(31.6%)
PGY-2	13(11.4%)	27(23.7%)	40(35.1%)
PGY-3	12(10.5%)	26(22.8%)	38(33.3%)
Gender			
Female	16(26.2%)	22(36.1%)	38(62.3%)
Male	7(11.5)	16(26.2%)	23(37.7%)
Age at PGY- 1 (years)			
<30	14(23.0%)	21(34.4%)	35(57.4%)
30–39	8(13.1%)	14(23.0%)	22(36.1%)
≥40	3(4.9%)	1(1.6%)	4(6.6%)
Race			
Asian	5(8.2%)	20(32.8%)	25(41.0%)
African American	6(9.8%)	10(16.4%)	16(26.2%)
Caucasian	4(6.6%)	6(9.8%)	10(16.4%)
Hispanic	4(6.6%)	6(9.8%)	10(16.4%)
Track			
Traditional	14(23.0%)	17(27.9%)	31(50.8%)
Underserved	9(14.8%)	21(34.4%)	30(49.2%)
Years elapsed since medical school graduation†			
<1	16(26.2%)	2(3.3%)	18(29.5%)
1–4	6(9.8%)	16(26.2%)	22(36.1%)
5–10	1(1.6%)	11(18.0%)	12(19.7%)
>10	9(14.8%)	0(0.0%)	9(14.8%)
Total	23(37.7%)	38(62.3%)	61(100%)

\*Residents in different postgraduate years; † $p<0.05$ .

graduation from medical school, there were no significant differences between US and non-US graduates.

Residents reported 8543 OB procedures in the procedure log and evaluation system between 1 July 2005 and 30 June 2008. Table 4 depicts significant variations in the number of reported procedures between residents by training year. Table 4 also shows that PGY-1 residents recorded the greatest number of OB procedures, with an average of 120 procedures, while PGY-2 and PGY-3 residents reported 85 and 30 OB procedures on average, respectively. Over the course of training, there was no significant difference in quantity of reported OB procedures between genders, with male residents recording

Table 4. Resident reported procedures by resident characteristics

Characteristic	Total procedure	Percent	Procedures/ Resident
PGY			
PGY-1(33)	4032	47.2	121.7±95.9
PGY-2(38)	3396	39.8	84.9±53.7
PGY-3(35)	1115	13.0	30.8±36.9
Gender			
Male	3341	39.1	167.1±106.0
Female	5202	60.9	140.6±145.6
Track			
Traditional	4782	56.0	170.8±157.8
Underserved	3761	44.0	129.7±101.6
US Graduate			
Non-US	6370	74.6	167.6±143.8
US	2173	25.4	114.4±101.0
Academic year*			
2005–2006	3304	37.4	89.3±83.0
2006–2007	3799	44.5	97.4±90.6
2007–2008	1440	18.0	42.4±32.1
Total	8543	100.0	95.2±102.7

PGY, Postgraduate year. \* $p<0.05$ .

167 OB procedures per year and female residents recording 140 OB procedures per year ( $p>0.05$ ). Non-US medical school graduates recorded a greater number of OB procedures than US medical school graduates. The residents enrolled in the traditional track reported a greater number of OB procedures than the residents enrolled in the underserved track.

Table 5 shows the procedure involvement, procedure category, and top 10 procedures reported by residents. Among all reported procedures, about 84% of the procedures were performed or independently performed by residents; 12.5% of procedures were not directly performed by residents, but the residents had an assisting role. Obstetric procedures, surgery procedures, and gynecological procedures were the top three categories reported. The most reported surgery procedures included incision and drainage-abscess, excision skin tag, simple laceration repair, and sebaceous cyst removal. These top three categories account for more than 75% of all reported procedures, while gynecology diagnosis and management procedures accounted for only 1.1%. The top ten reported





## Procedure evaluation of family medicine residents

Table 5. Procedure involvement, procedure category and top 10 procedures

Procedure	Frequency	Percent
Involvement*		
Assisted	1070	12.5
Independent	1054	12.3
Performed	6119	71.6
Supervised	300	3.5
Category		
GYN DX/MGMT	20	0.2
Behavioral Science	79	0.9
PED/NEO	138	1.6
PED/NEO DX/MGMT	201	2.4
OB DX/MGMT	238	2.8
Orthopedics	445	5.2
Medicine DX/MGMT	452	5.3
Medicine	538	6.3
GYN	572	6.7
Surgery	1141	13.4
OB	4719	55.2
Top 10 procedures		
Normal Vaginal Delivery	2386	27.9
Perineal Laceration Repair	620	7.3
Amniotomy	434	5.1
IUD Placement	260	3.0
Arthrocentesis Injection	235	2.8
Incision and Drainage - Abscess	232	2.7
IUPC Placement	207	2.5
Labor Induction	192	2.2
Fetal scalp monitor placement	158	1.8
Oxytocin for Argumentation of Labor	149	1.7

\*Procedure involvement:

Assisted: Guided another learner, with faculty supervision.

Independent: Performed, without faculty supervision.

Performed: Performed with faculty supervision.

Supervised: Guided another learner, without faculty supervision.

individual procedures were: normal vaginal delivery, perineal laceration repair, amniotomy, intrauterine device placement, incision and drainage – abscess, arthrocentesis injection, IUPC (Intrauterine Pressure Catheter) placement, labor induction, fetal scalp monitor placement, and oxytocin administration for augmentation of labor. Normal vaginal delivery made up 27.6% of all procedures.

Table 6 shows the OB procedure evaluation for residents by level of training based on three elements (competency, medical knowledge, and patient care). A total of 1263 procedures were evaluated by faculty; however, not all of the procedures included an evaluation of the three elements because of the involvement of another faculty member. The overall results showed that the residents' procedural skills, procedure-related medical knowledge, and patient care skills all improved with training. The procedural competency was assessed on a 1–6 scale, while knowledge and patient care were assessed on a 1–3 scale.

### OB procedure competency evaluation

The average scores for procedure competency at the PGY-1, -2, and -3 levels were 4.32, 5.11, and 5.41, respectively. There were significant differences in procedure competency as a function of PGY ( $p < 0.01$ ), indicating significant improvement of procedural skills during the training (Fig. 2). Significant differences also existed between the genders, with female residents performing better than male residents overall. The average procedure competency evaluation score for female PGY-1 residents was higher than male residents; the differences in procedure evaluation scores were less for PGY-2 residents. The evaluations of procedure competency for male residents surpassed female residents by PGY-3. Between PGY-1 and PGY-3, the procedure competency evaluation score for male residents increased 1.24 points compared to 0.12 points for female residents. The procedure competency evaluation scores were not significantly different in the overall evaluations between non-US medical school graduates and US medical school graduates, but US graduates scored higher on the procedural competency evaluations at PGY-1 and PGY-3 compared with the non-US graduate counterparts. Furthermore, non-US graduates showed greater improvement than US graduates. At PGY-1, the procedure competency evaluation score for underserved track residents was higher than residents in the traditional track. At PGY-2 and PGY-3 the procedure competency evaluation for the residents in both tracks was similar, although the residents in the traditional track had higher competency evaluation. The data analysis suggested that the improvement in procedural competency among different genders and tracks, and between US and non-US graduates varied over the years. Although there was no statistically significant



Table 6. Obstetrics procedures evaluation\* by postgraduate year, gender, medical school attended, and track

Characteristic	Procedure competency evaluation			
	PGY-1	PGY-2	PGY-3	Total
Gender				
Male	4.26±0.98(72)	5.08±0.74(396)	5.50±0.63(58)	5.01±0.83(526)
Female	5.25±0.50(4)	5.12±0.78(604)	5.37±0.68(129)	5.17±0.77(737)
US graduate				
Yes	5.00±0.00(3)	5.01±0.73(320)	5.47±0.62(49)	5.07±0.75(372)
No	4.29±0.99(73)	5.15±0.78(680)	5.39±0.68(138)	5.12±0.83(891)
Track				
Traditional	4.20±1.00(66)	5.12±0.74(475)	5.44±0.71(95)	5.07±0.83(636)
Underserved	5.10±0.32(10)	5.09±0.79(525)	5.38±0.61(187)	5.13±0.77(627)
Total	4.32±0.298(76)	5.11±0.77(1000)	5.41±0.66(187)	5.10±0.80(1263)
Characteristic	Procedure-related medical knowledge			
	PGY-1	PGY-2	PGY-3	Total
Gender				
Male	2.05±0.31(42)	2.38±0.51(257)	2.65±0.48(37)	2.37±0.51(336)
Female	2.00±0.00(1)	2.43±0.53(429)	2.58±0.50(73)	2.45±0.53(503)
US graduate				
Yes	3.00±0.58(1)	2.34±0.52(231)	2.64±0.49(22)	2.37±0.52(254)
No	2.02±0.27(42)	2.45±0.52(455)	2.59±0.49(88)	2.44±0.52(585)
Track				
Traditional	2.02±0.27(41)	2.44±0.52(345)	2.53±0.50(51)	2.41±0.52(437)
Underserved	2.50±0.71(2)	2.38±0.52(341)	2.66±0.48(59)	2.42±0.52(402)
Total	2.05±0.31(43)	2.41±0.52(686)	2.60±0.49(110)	2.42±0.52(839)
Characteristic	Patient care evaluation			
	PGY-1	PGY-2	PGY-3	Total
Gender				
Male	2.11±0.51(63)	2.36±0.52(356)	2.77±0.42(57)	2.38±0.53(476)
Female	2.50±0.58(4)	2.43±0.53(563)	2.64±0.50(123)	2.47±0.53(690)
US graduate				
Yes	2.67±0.58(3)	2.36±0.50(295)	2.69±0.51(48)	2.41±0.52(346)
No	2.11±0.51(64)	2.43±0.53(624)	2.68±0.47(132)	2.44±0.54(820)
Track				
Traditional	2.05±0.48(57)	2.43±0.50(422)	2.66±0.50(91)	2.43±0.52(570)
Underserved	2.60±0.52(10)	2.38±0.54(497)	2.71±0.46(89)	2.43±0.54(596)
Total	2.13±0.52(67)	2.41±0.52(919)	2.68±0.48(180)	2.43±0.53(1166)

PGY, Postgraduate year. \*Cronbach's alpha coefficient for the three criteria questions was 0.779. A reliability coefficient of 0.70 or higher is considered "acceptable" in most Social Science research situations.

difference in procedural competency between US and non-US graduates and different tracks, differences were noted between

the two groups during the training process. There were greater differences between different groups at PGY-1, but

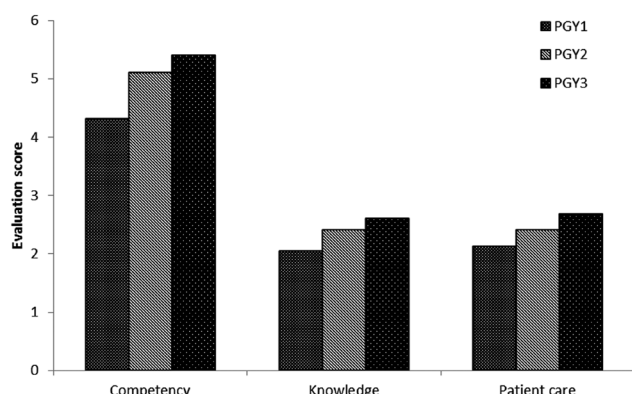


Fig. 2. Procedure evaluation for competency, knowledge, and patient care by postgraduate year\*.

\*Please note that competency is assessed on a 1–6 scale, while knowledge and patient care are assessed on a 1–3 scale.

the differences were less during PGY-3. A two-way ANOVA was used to examine the effect of residents' demographics and PGY levels on procedure performance evaluation scores. There was no significant interaction between the effects of gender, track, US versus non-US medical school, and PGY level on the procedure competency evaluation.

### OB medical procedure knowledge evaluation

The results in Table 6 indicate that the procedure-related medical knowledge level also improved during the training process. The average evaluation score for PGY-1, -2, and -3 residents was 2.05, 2.41, and 2.60, respectively ( $p < 0.001$ ). There were significant differences between the genders for procedure-related medical knowledge; the differences between tracks and US versus non-US medical school graduates diminished with training. The difference between different tracks at PGY-1 was 0.48, but decreased to 0.13 at PGY-3. Similarly, the differences between US and non-US graduates changed from 0.92 at PGY-1 to 0.05 at PGY-3. It should be noted that at PGY-2, female residents received higher evaluations on average, but at PGY-3 male residents received higher evaluations than females. The results of two-way ANOVA analysis on the effect of PGY and gender on medical knowledge indicate no interaction between genders and PGY on the procedure medical knowledge evaluation.

### OB patient care evaluation

The results indicate trends of improvement during the training process. The evaluation for PGY-1, -2, and -3 residents were 2.13, 2.41, and 2.68, respectively, and the differences across all PGY levels were significant. Generally, female and non-US medical graduate residents received better evaluations. By PGY-3, there were no significant differences between gender, US versus non-US graduates, and different tracks. Similar to the competency evaluation, the patient care evaluations for female residents at PGY-1 were better than the patient care evaluations for males, but by PGY-3 the evaluations of male residents surpassed the evaluations of female residents. Two-way ANOVA results indicate that gender, US versus non-US medical school, and track were not attributable to the patient care evaluation level, but the evaluation differences with respect to gender, track, and between US and non-US medical graduates during the training process from PGY-1 to PGY-3 were confirmed.

### Discussion

The efforts by the ACGME to standardize resident training and demonstrate objective clinical competency have led to more accurate documentation of resident competencies [12]. Procedure evaluation and certification by faculty is a necessary part of resident evaluation. Although family medicine residency programs require residents to log their procedures, few programs have efficient procedure evaluation and certification systems in place. Our procedure log system not only quantified the number of procedures and recorded detailed information about each procedure reported by residents, but it also qualified the residents' performance during the training process. This evaluation system not only encouraged our residents to log more procedures, but also the timeliness of procedure reporting was improved.

There are several procedural evaluation tools available [13]; however, there is no gold standard about how to best assess resident procedural competence. The assessment tools can be divided into two different types (generic assessment scales and procedure-specific skills assessment scales) [14]. Because family medicine residents are trained to manage a wide variety of adult medical conditions and perform the associated procedures, it is not feasible to have specific-procedure





evaluations for all the procedures within the scope of family practice. Furthermore, Martin [15] reported that global ratings are a better method of assessing a resident than a task-specific checklist. We used the generic approach by designing a simple procedure evaluation system with only three questions, a system which is easy to understand and implement.

Our data showed there were significant variations in the number of logged OB procedures by gender, residency track, and US versus non-US medical graduates, in part because residents stopped reporting procedures after the procedure was certified. This wide inconsistency stresses the importance of tracking and recording procedures by residents, and highlights the need for implementation of efficient and innovative measures to ensure residents not only log their procedures in a timely manner, but also involve certification by a faculty member.

Residents who performed OB procedures were evaluated for competency, medical knowledge, and patient care during the procedure. These three observable criteria that were used to evaluate procedure performance are simple, yet comprehensive enough to make the procedure evaluation program feasible. Moreover, a procedure evaluation also aligns with the most current ACGME accreditation system for residency training. In addition, the reliability analysis results supported the conclusion that the questionnaire led to better internal consistency and reliability for the three evaluation criteria (Cronbach's  $\alpha=0.779$ ). Cronbach's  $\alpha$  demonstrates how well a set of items or variables measures a single unidimensional latent construct. Cronbach's  $\alpha$  can reflect internal validity of a questionnaire, in this case the residents' procedural competency [16]. A reliability coefficient of 0.70 or higher is considered "acceptable" in most social science research situations [17].

The evaluation results show that residents' procedural performance improved during training in competency, procedure-related medical knowledge, and patient care skills. Generally, there were differences in residents' procedure evaluations with respect to gender, tracks, and US versus non-US graduates, but the differences decreased from PGY-1 to PGY-3. It is interesting to note that at the start of residency, female residents received higher procedure evaluations than their male counterparts, but by PGY-3 the procedure performances

of male residents surpassed the procedure performances of female residents. These gender differences in resident performance have been previously reported, especially related to the level of confidence during clinical evaluations [18, 19].

Finally, because procedures are documented in an electronic medical record system, it will be possible to integrate procedure evaluation with the medical records system in the future to facilitate resident logging procedures and improve data quality, accuracy, and timeliness of procedure evaluations.

### Limitation

This procedure evaluation system was designed by our department faculty members, and the validity was not evaluated. This system was a single family medicine residency program experience and has not been tested in other residency programs or other institutions. There were significant variations in reporting procedures, which affected the results and conclusions. Given that this is a procedure competency evaluation, extra post-graduate training among residents who had a hiatus of more than 10 years between graduation and the data collection period (14.8%) could be a potential confounder. Unfortunately, information about additional training, such as previous OB/GYN experience, was not collected among participants. The procedures evaluated in this system only included OB procedures, and all other family medicine resident required procedures should be evaluated and certified.

### Conflict of interest

The authors declare no conflict of interest.

### Funding

Residency Training Grant

Technical Title of Project: Residency Training in Primary Care, 1 D22 HP 00223-04.

Funding Agency: Department of Health and Human Services (Public Health Service), Health Resources and Services Administration, Bureau of Health Professions.

Investigator Relationship: Corboy, JE, Principal Investigator; Project Director.

Dates of Funding: July 1, 2004–June 30, 2007.

Total Direct Cost: = \$534,163.



## Procedure evaluation of family medicine residents

### References

1. Wigton RS. Measuring procedural skills. *Ann Intern Med* 1996;125(12):1003–4.
2. Norris TE, Cullison SW, Fihn SD. Teaching procedural skills. *J Gen Intern Med* 1997;12(Suppl 2):S64–70.
3. Kovacs G. Procedural skills in medicine: linking theory to practice. *J Emerg Med* 1997;15(3):387–91.
4. Dickson GM, Chesser AK, Keene Woods N, Krug NR, Kellerman RD. Self-reported ability to perform procedures: a comparison of allopathic and international medical school graduates. *J Am Board Fam Med* 2013;26(1):28–34.
5. Cheung BM, Ong KL, Cherny SS, Sham PC, Tso AW, Lam KS. Diabetes prevalence and therapeutic target achievement in the United States, 1999 to 2006. *Am J Med* 2009;122(5):443–53.
6. Nasca TJ, Philibert I, Brigham T, Flynn TC. The next GME accreditation system – rationale and benefits. *N Engl J Med* 2012;366(11):1051–6.
7. Nothnagle M, Sicilia JM, Forman S, Fish J, Ellert W, Gebhard R, et al. Required procedural training in family medicine residency: a consensus statement. *Fam Med* 2008;40(4):248–52.
8. Wetmore SJ, Rivet C, Tepper J, Tatemichi S, Donoff M, Rainsberry P. Defining core procedure skills for Canadian family medicine training. *Can Fam Physician* 2005;51:1364–5.
9. Wigton RS, Blank LL, Nicolas JA, Tape TG. Procedural skills training in internal medicine residencies. A survey of program directors. *Ann Intern Med* 1989;111(11):932–8.
10. Hawes R, Lehman GA, Hast J, O'Connor KW, Crabb DW, Lui A, et al. Training resident physicians in fiberoptic sigmoidoscopy. How many supervised examinations are required to achieve competence? *Am J Med* 1986;80(3):465–70.
11. Cuda AS. Mixed methods curriculum evaluation: maternity care competence. University of Washington; 2013.
12. Bhalla VK, Bolduc A, Lewis F, Nesmith E, Hogan C, Edmunds JS, et al. Verification of resident bedside-procedure competency by intensive care nursing staff. *J Trauma Nurs* 2014;21(2):57–60; quiz 61–2.
13. Jelovsek JE, Kow N, Diwadkar GB. Tools for the direct observation and assessment of psychomotor skills in medical trainees: a systematic review. *Med Educ* 2013;47(7):650–73.
14. Ahmed K, Miskovic D, Darzi A, Athanasiou T, Hanna GB. Observational tools for assessment of procedural skills: a systematic review. *Am J Surg* 2011;202(4):469–80.e6.
15. Martin JA, Regehr G, Reznick R, MacRae H, Murnaghan J, Hutchison C, et al. Objective structured assessment of technical skill (OSATS) for surgical residents. *Br J Surg* 1997;84(2):273–8.
16. Ash DS; American Society of Health-System Pharmacists. ASHP's management pearls. 2009, Bethesda, Md.: American Society of Health-System Pharmacists; 107 p.
17. Tavakol M, Dennick R. Making sense of Cronbach's alpha. *Int J Med Educ* 2011;2:53–5.
18. Craig LB, Smith C, Crow SM, Driver W, Wallace M, Thompson BM. Obstetrics and gynecology clerkship for males and females: similar curriculum, different outcomes? *Med Educ Online* 2013;18:21506.
19. Bartels C, Goetz S, Ward E, Carnes M. Internal medicine residents' perceived ability to direct patient care: impact of gender and experience. *J Womens Health (Larchmt)* 2008;17(10):1615–21.