



Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.

Original Article

A core competence-based objective structured clinical examination (OSCE) in evaluation of clinical performance of postgraduate year-1 (PGY₁) residents

Ying-Ying Yang^{a,b,d}, Fa-Yauh Lee^{a,b,d,*}, Hui-Chi Hsu^{a,b,d}, Chin-Chou Huang^{a,c,d},
Jaw-Wen Chen^{a,c,d}, Wen-Shin Lee^{a,b,d}, Chiao-Lin Chuang^{a,b,d}, Ching-Chih Chang^{a,b,d},
Hao-Min Chen^{a,c,d}, Chia-Chang Huang^{a,c,d}

^a Division of General Medicine, Taipei Veterans General Hospital, Taipei, Taiwan, ROC

^b Department of Medicine, Taipei Veterans General Hospital, Taipei, Taiwan, ROC

^c Department of Medical Research and Education, Taipei Veterans General Hospital, Taipei, Taiwan, ROC

^d Department of Medicine, National Yang-Ming University School of Medicine, Taipei, Taiwan, ROC

Received October 19, 2010; accepted January 31, 2011

Abstract

Background: Clinical competency certifications are important parts of internal medicine residency training. This study aims to evaluate a composite objective structured clinical examination (OSCE) that assesses postgraduate year-1 (PGY₁) residents' acquisition of the six core competencies defined by the Accreditation council for Graduate Medical Education (ACGME).

Methods: Six-core-competency-based OSCE was used as examination of the clinical performance of 192 PGY₁ residents during their 3-month internal medicine training between 2007 January and 2009 December. For each year, the reliability of the entire examination was calculated with Cronbach's alpha.

Results: The reliability of six-core-competency-based OSCE was acceptable, ranging from 0.69 to 0.87 between 2007 and 2009. In comparison with baseline scores, the summary scores and core-competency subscores all showed significant increase after PGY₁ residents finished their 3-month internal medicine training program.

Conclusion: By using a structured development process, the authors were able to create reliable evaluation items for determining PGY₁ residents' acquisition of the ACGME core competencies.

Copyright © 2011 Elsevier Taiwan LLC and the Chinese Medical Association. All rights reserved.

Keywords: Objective structured clinical examination; Postgraduate year-1 residents; Six core competencies

1. Introduction

Competency assessment has been an obligation and an ongoing challenge for those institutions responsible for the training and certification of physicians. Such assessment has had the practical function of establishing minimal professional

standards that ensure the basic fitness of future physicians. The introduction of in-training assessment using clinicians to assess the performance of residents is a task. The assessment method must provide reliable results.

The primary goals of internal medicine residency training program are to produce competent practitioners. In comparison with other conventional assessment methods, the Objective Structured Clinical Examination (OSCE) is a multidimensional practical examination of clinical performance of residents. Initially, experience with the OSCE has been somewhat limited in the United States, Canada, United Kingdom and Australia.^{1–3}

* Corresponding author. Dr. Fa-Yauh Lee, Division of General Medicine, Department of Medicine, Taipei Veterans General Hospital, 201, Section 2, Shih-Pai Road, Taipei 112, Taiwan, ROC.

E-mail address: fylee@vghtpe.gov.tw (F.-Y. Lee).

However, this evaluation method has emerged elsewhere, in Asia, including Japan, Korea and Taiwan, as the premiere method for assessing clinical competence and uncovered some competency deficits that are missed by other methods.

Nevertheless, wide variation exists in the scoring schedules of OSCE. Recently, experts suggested that a more constructed core competency-based modification of the scoring system of OSCE is mandatory to appropriate evolution of clinical performance of residents.^{4–7}

In the present study, we had three goals: (1) to determine the reliability of the six-core-competency-based OSCE in sequential testing of PGY₁ residents; (2) to compare core-competency acquisition before and after the internal medicine training program of PGY₁ residents; and (3) to determine the usefulness of the information gained from PGY₁ resident clinical performance.

2. Methods

2.1. PGY₁ program

The outbreak of the severe acute respiratory syndrome (SARS) epidemic that occurred during 2003 exposed serious deficiencies in Taiwan's medical care and public health care systems, as well as its medical education system. The Department of Health, Executive Yuan of Taiwan ROC has

made efforts in promoting the “Project of Reforming Taiwan's Medical Care and Public Healthcare System” since the spread of SARS was controlled. The reform of the medical care system aimed to provide better holistic medical treatment to people. The strategies and methods are strengthening the improvement of resident education and quality of medical care. A project titled “Postgraduate General Medical Training Program” was announced by the Department of Health in August 2003. Through this project, each doctor in his/her first year of residency (including internal medicine, family medicine, surgery, pediatrics, dermatologist, ophthalmologist, etc.) is required to fulfill 3 months of an internal medical training course, along with 36 hours of basic courses. In the past, there was no such program in Taiwan to provide general medical training for medical students after graduation. Therefore, the goal of this program is to ensure that all PGY₁ residents have acquired Accreditation Council for Graduate Medical Education (ACGME) core competence in internal medical care.

2.2. OSCE setting

A comprehensive 6-station OSCE was administered to 192 PGY₁ residents between 2007 January and 2009 December at Taipei Veterans General Hospital (Taipei VGH). Taipei VGH is a regional medical center that provides primary and tertiary care to active duty and retired military members, and their

Table 1
The content of objective structured clinical examination (OSCE) stations of postgraduate year-1 (PGY₁) residents

Name of OSCE station	2007	2008	2009	Aspects of competence/skills assessed
Use of clinical information	○	○	○	PC- interviewing; physical examination MK- investigatory and analytic thinking
Organization and orderliness	○	○	○	P- sensitiviting to cultural, age, gender, and disability issues ICS- creation of therapeutic relations with patients
Patient safety and ethical issues	●	●		SBP- understanding interaction of their practice with the larger system; advocating for patient within the health care system PBLI- analyzing own practice for needed improvement
Creation of therapeutic relations with patients		●	●	SBP- knowledge of practice and delivery system PBLI- using evidence from scientific studies (EMB); using of information technology
Providing patient-centered care	●			P- respect and altruism ICS- listening
Counseling and educating patients and family	●	●		MK- knowledge and application of basic science PC- couesling and educate patients and families; preventive health service; informed decision-making
Decision-making (clinical judgment)			●	SBP- understanding interaction of their practice with the larger system; advocating for patient within the health care system PBLI- analyzing own practice for needed improvement
Clinical differential diagnosis			●	MK- knowledge and application of basic science PC- counseling and educate patients and families; preventive health service; informed decision-making
Improvement of quality of clinical care		●		ICS- listening P- respect and altruism
Using evidence-based practice	○			PBLI- use evidence from scientific studies (EMB); use of information technology SBP- knowledge of practice and delivery system
Interaction with whole medical system			●	ICS- listening P- respect and altruism
Reliability (Cronbach's alpha) of OSCE	0.72	0.69	0.87	

○ = the station implemented for the year; ● = the station implemented and standardized. Patients used for the year; ICS = interpersonal and communication; MK = medical knowledge; P = professionalism; PBLI = problem-based learning and improvement; PC = patient care; SBP = systems-based practice; each OSCE station assessed two domains of ACGME core competencies; each domain (skill) of ACGME competency was equally assessed twice in 6-station OSCE of every year.

dependents. Taipei VGH serves as the primary teaching hospital for its internal medicine residency program. The examination was conducted on two consecutive Friday afternoons during the internal medicine training program of all PGY₁ residents.

The OSCE consisted of 6 clinical problems (Table 1); each clinical problem consists of six core competencies defined by the ACGME (Table 2). The content of each clinical problem is included in Table 1.⁸ The OSCE has neither a written component nor technical skills stations, but is entirely performance-based. At some stations, standardized patients were used to mimic the clinical problems of actual patients. A faculty rater graded each PGY₁ resident according to a given set of 12 predetermined items (using 2 items to evaluate each aspect of the ACGME core competencies) presented in the form of a checklist. The checklists for all 6 OSCE stations have similar items to cover ACGME core competency. The items that PGY₁ presented completely were scored as “1” and those not presented were scored as “0”. All faculty raters attended serial training sessions that included extensive instructions on how to use the checklist in practice rating sessions. At each OSCE station, the raters acted as passive evaluators and were instructed not to guide or prompt the PGY₁ residents.

At each station, the summary scores were the sum of all the checklist items, and the six core competency subscores were the sum of specific items for each competency. When presented, all scores were translated into 100 percentages. Finally, each PGY₁ resident got summary scores and core competence subscores of the 6 OSCE stations. The “borderline group

method” was used to set the standard of “pass”. Each station’s “pass” score was the mean of the scores of PGY₁ residents whose OSCE scores were rated “borderline”.^{9,10}

2.3. Analyzing the reliability of the scoring system

Preliminary studies with senior residents permitted us to identify and eliminate unreliable and ambiguous items, resulting in an instrument consisting of 12 items, with 2 items assessing each of the six core competencies (Table 2).

To evaluate the reliability of the OSCE checklist as measure of their competency, we conducted a factor analysis of the scoring system (items) of OSCE. Beginning with a principal component analysis (PCA),^{11,12} we identified six components with eigen values greater than 1.0, suggesting six factors underlying the score of PGY₁ OSCE (Table 2). A correlation table of scores for each station was calculated, and used to exclude stations that correlated poorly with all other stations from further analysis.

2.4. Study protocols and reliability (internal and external)

In 2007, we tested 60 PGY₁ residents; in 2008, 72 and in 2009, 60. Furthermore, we assessed internal and test/retest (external) reliability of the six-core-competency-based OSCE by computing Cronbach’s alpha coefficients.¹³ Inter-rater reliability was assessed by percent agreement and kappa statistical analysis. In a subgroup of PGY₁ residents ($n = 70$), the same clinical problems with identical OSCE setting were

Table 2
Factor analysis of six core competence-based checklists in objective structured clinical examination (OSCE)

Domain of ACGME competence and skills	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6
Patient care skills						
Interviewing	0.40					
Counseling and educating patients and families	0.43					
Physical examination	0.43					
Preventive health service	0.55				0.31	
Informed decision making	0.34					
Interpersonal and communication skills						
Creation of therapeutic relations with patients		0.87				
Listening		0.65				
Professionalism skills						
Respect and altruism			0.59			
Sensitivity to cultural, age, gender, and disability issues			0.80			
Practice-based learning and improvement skills						
Analyzing own practice for needed improvement				0.73		
Using evidence from scientific studies (EMB)				0.68		0.41
Using information technology				0.71		
Systems-based practice skills						
Understanding interaction of their practice with the larger system					0.57	
Advocate for patient within the health care system					0.68	
Knowledge of practice and delivery system					0.74	
Medical knowledge skills						
Investigatory and analytic thinking						0.62
Knowledge and application of basic science						0.71

Data are eigen values (0–1); correlation coefficients between items assessing specific skills of each ACGME competency with each factor.

used to test their clinical performance before and after finishing the 3-month internal medicine training program. For this group of PGY₁ residents, the difference in their pre-course and post-course OSCE scores mainly came from the sequential time-point of assessment. Actually, the raters and standardized patients were unchanged between pre-course and post-course OSCE. Accordingly, the Cronbach's alpha coefficients of test/retest (external) reliability reflect the stability of the OSCE.

2.5. Data analysis

Data were shown as mean \pm standard deviation (SD). Statistical analysis was performed using SPSS version 13.0 (SPSS Inc., Chicago, IL, USA). Analysis of variance was used to compare the means between summary OSCE scores and core-competency subscores, McNemar and Chi-square test were used to analyze the subscores and pass rates between different summary scores and core-competency subscores. Paired 2-samples *t* test was used to compare the performance of OSCE between pre-and post-course of 3-month PGY₁ internal medicine training period. Inter-rater reliability was analyzed with Kappa statistics. Factor analysis was assessed with PCA. Spearman-Brown prophecy formula was used to calculate the number of OSCE stations needed to reach the desired level of reliability (0.8). An α of <0.05 was accepted as statistically significant.

3. Results

The internal reliability of each year's six-core-competency-based OSCE had excellent agreement (Table 1, Cronbach's alpha in 2007 = 0.728, in 2008 = 0.69, in 2009 = 0.87). The overall 3-year internal reliability of summary OSCE scores and core-competency subscores were also acceptable (Table 3). In addition to professionalism aspects of core-competency, the overall 3-year inter-rater reliability (Kappa statistics) on the summary OSCE scores and core-competency subscores was also good (Table 4).

In Fig. 1A, a significant difference in the performance between different aspects of core competency ($F = 49.8$, $p < 0.05$) was noted. PGY₁ residents had the highest pass rate in the aspect of interpersonal and communication skills (81%), whereas the lowest pass rate was noted in the aspect of professionalism (53%). Interestingly, the corresponding data

Table 4
Inter-rater reliability of OSCE scores derived from PGY₁ residents

	Percent agreement	Kappa
Average of total summary OSCE score	0.91	0.732
Patient care subscore	0.93	0.71
Medical knowledge subscore	0.88	0.74
Practice-based learning and improvement subscore	0.90	0.77
Interpersonal and communication skills subscore	0.83	0.69
Professionalism subscore	0.89	0.79
Systems-based practice subscore	0.92	0.68

Two raters for each OSCE station; ($n = 192$).

in Table 3 confirmed that the highest subscores were noted in the aspects of interpersonal and communication skills, whereas the lowest subscores were observed in the professionalism aspect.

The eigen values (with variance accounted for presented in parentheses) of the six factors resulting from this PCA of our study were 6.28 (28%), 2.93 (39%), 2.17 (47%), 2.25 (54%), 1.4 (59%) and 1.4 (63%), respectively. Additionally, the resulting scree plot of the eigen values revealed that the leveling off to a straight horizontal line occurred after the sixth eigen value. Then, we performed a principal axis factor analysis with oblique rotation to a preselected six-factor solution. The factors that emerged from this analysis clearly reflect each of the six core competencies of residency (Table 2). Intriguingly, items from each competency loaded on their intended factor and did not load with items from different scales of competencies. In other words, the factor analysis resulted in a nearly perfectly clean structure, providing evidence that our theoretically derived items in checklist of PGY₁ OSCE were distinct and evident in the reflection of core competence of their clinical performance.

In a sub-group of PGY₁ residents who received both the pre-course and post-course six core-competency-based OSCE, the test/retest (external) reliability was borderline (Table 4). However, we clearly observed significantly increased OSCE scores and subscores after the internal medicine training program. Meanwhile, the pass rates were markedly improved in all aspects of core competencies in PGY₁ residents (Fig. 1B). The most improved aspects were patient care and professionalism (Table 5). In contrast, the performance in systems-based practice was not significantly improved after the training program.

Table 3
Reliability (internal consistency) of OSCE scores derived from PGY₁ residents

	^a Score (%)	Range (%)	Reliability: (Cronbach's alpha)
Average of total summary OSCE score	74.9 \pm 12.1	41.8–81.9	0.73
Patient care subscore	83.1 \pm 17.9	49.9–87.1	0.673
Medical knowledge subscore	76.8 \pm 14.9	50.1–82.9	0.672
Practice-based learning and improvement subscore	82.1 \pm 11.2	61.9–82.1	0.642
Interpersonal and communication skills subscore	84.1 \pm 12.9	54.2–88.1	0.709
Professionalism subscore	70.2 \pm 12.7	44.7–78.1	0.683
Systems-based practice subscore	76.5 \pm 19.7	34.7–80.1	0.702

^a Score = (obtained points/maximum possible points) \times 100%; ($n = 192$). Data were expressed as mean \pm SD.

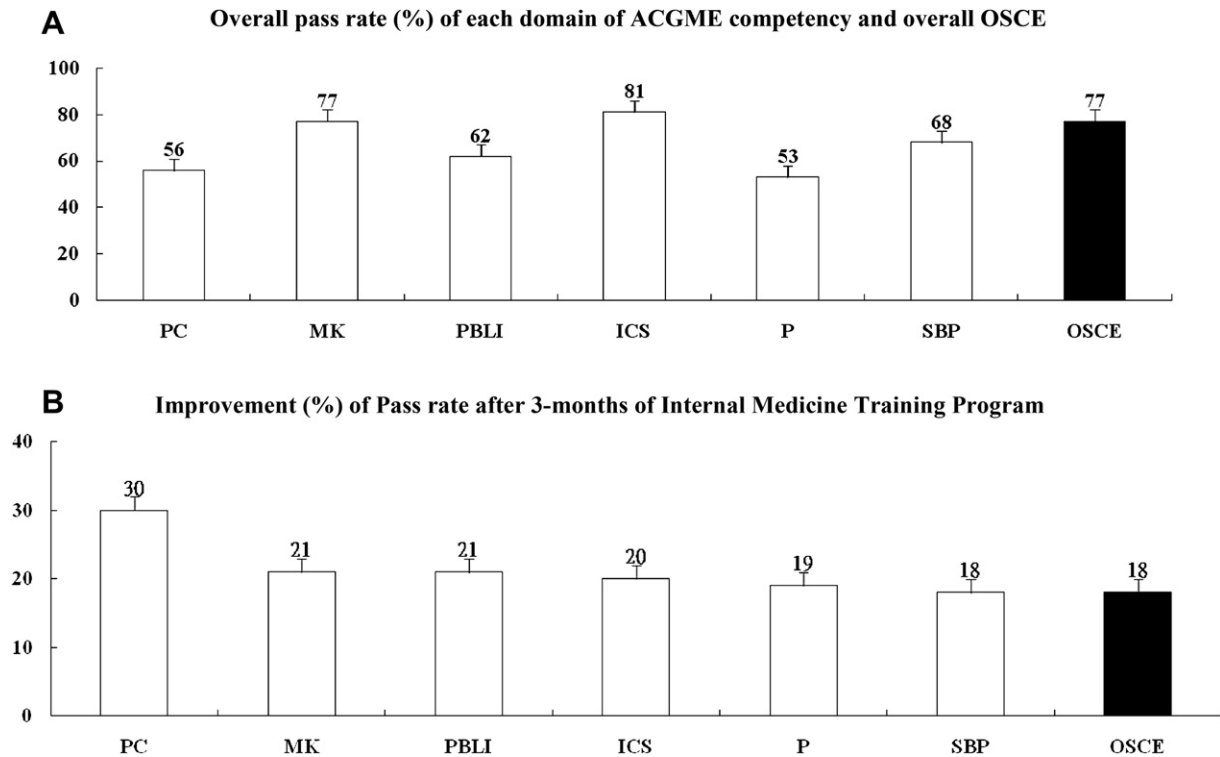


Fig. 1. (A) The overall pass rate (pass students/total students \times 100%) and (B) the degree of the improvement of core-competency-based PGY₁ OSCE. PC: patient care; MK: medical knowledge; PBLI: practice-based learning and improvement; ICS: interpersonal and communication skills; P: professionalism; SBP: systems-based practice.

4. Discussion

In this study, a composite global performance score was determined for each core competency from merging separate items. This is in contrast to unclassified global assessment OSCE scores that may not provide an accurate assessment of the PGY₁ resident performance.^{14,15} Meanwhile, our items were developed by experienced medical education leaders and were based largely on the conceptual framework contained in the model of the clinical practice of internal medicine.¹⁶ It had been reported that using structured scoring guidelines can appropriately assess whether PGY₁ residents have acquired core competency during training.¹⁷

In Fig. 1B and Table 4, higher scores of core-competency-based OSCE than baseline were found after the training program. This result supported the discriminatory reliability of

the six core competency evaluation process of PGY₁ OSCE in our study.

PCA is a form of factor analysis. In our study, factor-analytic techniques were used to detect structure in the relationship between variables. With PCA, a “line” (an eigen vector) is defined in the scatter plot that illustrates maximal variance. This is also known as the “factor proportion” or the “factor load”. The square of the factor proportion gives the percentage of the variability that an element contributes to a system. After this first factor has been extracted, the next line (second eigen vector) that maximizes variances is defined. In this way, one is able to determine the contribution of each consecutive factor to the overall variability of the results.^{11,12}

Reliability on the aspects of professionalism and system-based practice were lower (Cronbach’s alpha = 0.31 and 0.419, respectively) than the other four aspects of core

Table 5
Comparison of performance of PGY₁ residents between pre-course and post-course of a 3-month internal medicine training program

	Pre-course	Post-course	Reliability: (Cronbach’s alpha)
Average of total summary OSCE score	74.9 \pm 12.3	79.9 \pm 17.8*	0.602
Patient care subscore	73.3 \pm 15.2	89.5 \pm 13.9**	0.631
Medical knowledge subscore	75.9 \pm 10.1	90.1 \pm 16.3**	0.521
Practice-based learning and improvement subscore	72.3 \pm 12.6	81.4 \pm 20.2*	0.564
Interpersonal and communication skills subscore	83.2 \pm 21.5	84.0 \pm 18.3	0.477
Professionalism subscore	69.4 \pm 10.9	72.9 \pm 21.2**	0.499
Systems-based practice subscore	70.9 \pm 8.2	82.1 \pm 12.7*	0.621

* $p < 0.05$ versus pre-course.

** $p < 0.01$ (paired t test); ($n = 70$).

Data were expressed as mean \pm SD; Score = (obtained points/maximum possible points) \times 100%.

competency (Cronbach's alpha: patient care = 0.577; medical knowledge = 0.74; interpersonal and communication skills = 0.576; problem-based learning and improvement = 0.596). Certainly, the generally low inter-rater reliability of core-competency-based OSCE indicated that raters were still not familiar with the scoring system in our study. In other words, program directors should further unify the consensus between raters about the core-competency scoring system.¹⁷ Actually, it has been suggested that OSCE is not a good tool to assess medical knowledge. Thus, it is essential for program directors to add chart stimulation recall, oral and multiple choice question written examinations to assess medical knowledge in the future.^{14–16}

In our study, PGY₁ residents should complete each station within 15 minutes, and a 5-minute changeover time was allowed between stations; total testing time was 2 hours. It had been established that limited testing time is another possible factor for the low inter-rater reliability. Indeed, it was difficult for raters to judge 12 items on the checklist during every 15-minute station in our study. The literature is clear that longer periods improve the reliability of OSCE.⁶ Accordingly, longer test period is mandatory for more widely focused complex clinical problem in our core-competency-based OSCE.

The significant improvement in performance on patient care was accompanied with an acceptable reliability (Cronbach's alpha = 0.631) after finishing the internal medicine training program. Nonetheless, the insignificant improvement of professionalism was associated with relatively good reliability (Cronbach's alpha = 0.499). The inconsistent findings in this study suggested that the professionalism competency in the checklist should be adjusted in the future. However, our study still had the following limitations. First, scores in each group have a major impact on the mean scores and standard deviations at each year level. Despite this limitation, relatively good test/retest reliability and a significant increase in scores with time spent in the training program support the reliability of core-competency-based OSCE in our study.

Second, only six examination stations were included in our OSCE; previous studies had reported that a greater number of stations would increase the reliability. If this examination were to be used for high-stakes purposes, such as certification, a reliability of at least 0.8 would be desirable. One strategy for increasing the reliability over the present level of about 0.65 is to increase the number of stations. Using the Spearman-Brown formula, it is estimated that two additional stations would be necessary using the bench model with global scoring to achieve a reliability of 0.8. In the present study, each OSCE station measured two aspects of ACGME core competency and assessed two to four specific skills. Thus, another solution to increase reliability of ACGME competency-based OSCE might be to increase the number of OSCE stations and evaluate only one specific skill for each aspect of ACGME competency in the future.

Third, variation of PGY₁ residents' OSCEs because of rater bias is a possibility. Faculty raters may have previous experience

with a PGY₁ resident which could introduce positive or negative bias in scoring.

Fourth, several of the competencies as defined by ACGME were found in more than one of the six major categories. For example, the concept of working with other health care professionals is incorporated in patient care, interpersonal and communication skills, and systems-based practice. In real practice, a single question for this was listed under system-based practice: "works effectively as a member or leader of the team to provide patient-focused care: understands how his or her actions affect others". It could arguably be placed in either of the other two competencies. Actually, each of the items to be rated was intended to reflect a specific type of observable skill by the resident, not a personality trait.

Taken together, our study was characterized by a longitudinal experience with the core competency-based OSCE to evaluate internal medicine training program of PGY₁ residents over 3 years. In line with previous studies, our results indicate a number of useful findings for program directors.^{18,19} Notably, the overall internal reliability of our core competency-based OSCE progressively improved within 3 years (from 0.72 to 0.87). Furthermore, the internal reliability between the six core competencies was evenly distributed between 0.642 and 0.709. In contrast, the inter-rater reliability between different aspects of core competency was quite variable.

Acknowledgments

We would like to thank staff members of our division and department for their efforts in production and preparation of the assessment, and the raters, who participated in and helped improve the course and assisted in data collection.

References

1. Cohen R, Reznick RK, Taylor BR, Provan J, Rothman A. Reliability and validity of an objective structured clinical examination in assessing surgical residents. *Am J Surg* 1990;**160**:302–5.
2. Petrusa ER, Blackwell TA, Ainsworth MA. Reliability and validity of an objective structured clinical examination for assessing the clinical performance of residents. *Arch Intern Med* 1990;**150**:573–7.
3. Reznick R, Smee S, Rothman A, Chalmers A, Swanson D, Dufresne L, et al. An objective structured clinical examination for the licentiate: report of the pilot project of the medical council of Canada. *Acad Med* 1992;**67**:487–94.
4. Grand' Maison P, Lescop J, Rainsberry P, Brailovosky CA. Large scale use of an objective structured clinical examination for licensing family physicians. *Can Med Assoc J* 1992;**146**:1735–40.
5. Matsell DG, Wolfish NM, Hsu E. Reliability and validity of the objective structured clinical examination in pediatrics. *Med Educ* 1991;**25**:293–9.
6. Newble DI, Swanson DB. Psychometric characteristics of the objective structured clinical examination. *Med Educ* 1998;**22**:325–34.
7. Roberts J, Norman G. Reliability and learning from the objective structured clinical examination. *Med Educ* 1990;**24**:219–23.
8. Leach DC. The ACGME competencies: substance or form? *J Am Coll Surg* 2001;**192**:396–8.
9. Murto-Humphrey S, Mac Fadyen JC. Standard setting: a comparison of case-author and modified borderline-group methods in a small-scale OSCE. *Acad Med* 2002;**77**:729–32.

10. Boursiscot KAM, Roberts TE, Pell G. Using borderline methods to compare passing standards for OSCEs at graduation across three medical schools. *Med Educ* 2007;**41**:1024–31.
11. Chesser AM, Laing MR, Miedzybrodzka ZH, Brittenden J, Heys SD. Factor analysis can be a useful standard-setting tool in a high-stakes OSCE assessment. *Med Educ* 2004;**38**:825–31.
12. Gorsuch R. *Factor analysis*. 2nd ed. New Jersey, USA: Lawrence Erlbaum associates; 1983. p. 3–5.
13. Frey K, Edwards F, Alftman K, Spahr N, Gorman S. The “Collaborative Care” curriculum: an educational model addressing key ACGME core competencies in primary care residency training. *Med Educ* 2003;**37**: 786–9.
14. Swing SR. Assessing the ACGME general competencies: general consideration and assessment methods. *Acad Emerg Med* 2002;**9**: 1278–88.
15. Reisdorff EJ, Carlson DJ, Reeves MM, Walker G, Hayes QM, Reynolds B. Quantitative validation of a general competency composite assessment evaluation. *Acad Emerg Med* 2004;**11**:881–4.
16. Holmboe ES, Hawkins RE. Methods for evaluating the clinical competence of residents in internal medicine: a review. *Ann Intern Med* 1998; **129**:42–8.
17. Sloan DA, Donnelly MB, Schwartz RW, Strodel WE. The objective structured clinical examination: the new gold standard for evaluating postgraduate clinical performance. *Ann Surg* 1995;**222**:735–42.
18. Schwartz RW, Wizke DB, Donnelly MB, Stratton T, Blue AV, Slovan DA. Assessing residents’ clinical performance: cumulative results of a four-year study with the objective structural clinical examination. *Surgery* 1998;**124**:307–12.
19. Saunder MJ, Yeh CK, Hou LT, Katz MS. Geriatric medical education and training in the United States. *J Chin Med Assoc* 2005;**68**:547–56.